



## 

(43) International Publication Date 12 September 2002 (12.09.2002)

**PCT** 

# (10) International Publication Number WO 02/070703 A2

(51) International Patent Classification<sup>7</sup>: C12N 15/12, C07K 14/435, C12N 5/10, A0111 5/00, C07K 16/18, C12Q 1/68

Lucia, QLD 4072 (AU). **HOEGH-GULDBERG, Ian, Ove** [AU/AU]; 73 Seventh Avenue, St. Lucia, QLD 4072 (AU). **PRESCOTT, Mark** [AU/AU]; 49 Argyll Street, Malvern East, VIC 3145 (AU).

- (21) International Application Number: PCT/GB02/00928
  - .
- (22) International Filing Date: 1 March 2002 (01.03.2002)
- (25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

60/273,227 2 March 2001 (02.03.2001) US PR 3874 21 March 2001 (21.03.2001) AU 60/329,816 15 October 2001 (15.10.2001) US

(71) Applicants (for all designated States except US): NU-FARM LIMITED [AU/AU]; 103-105 Pipe Road, Laverton, VIC 3028 (AU). THE UNIVERSITY OF QUEENSLAND [AU/AU]; St. Lucia, QLD 4072 (AU).

(71) Applicant and

- (72) Inventor (for GB only): JONES, Elizabeth, Louise [GB/GB]; Frank B. Dehn & Co., 179 Queen Victoria Street, London EC4V 4EL (GB).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): KARAN, Mirko [AU/AU]; 2 Balfour Close, Watsonia, VIC 3087 (AU). BRUGLIERA, Filippa [AU/AU]; 11 Kalimna Street, Preston, VIC 3072 (AU). MASON, John [AU/AU]; 9/999 Rathdowne Street, Carlton North, VIC 3054 (AU). DOVE, Sophie, Gwendoline [AU/AU]; 73 Seventh Avenue, St.

- (74) Agents: JONES, Elizabeth, Louise et al.; Frank B. Dehn & Co., 179 Queen Victoria Street, London EC4V 4EL (GB).
- (81) Designated States (national): AE, AG, AL, AM, AT (utility model), AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ (utility model), DE (utility model), DK (utility model), DM, DZ, EC, EE (utility model), ES, FI (utility model), GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK (utility model), SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.
- (84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

#### Published:

 without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: CELL VISUAL CHARACTERISTIC-MODIFYING SEQUENCES

(57) Abstract: The present invention relates generally to peptides, polypeptides or proteins having one or more amino acids or one or more amino acid sequences which exhibit color-facilitating properties, either on their own or following interaction with one or more other amino acids and to nucleic acid molecules encoding same. Such peptides, polypeptides and proteins are referred to herein as "color-facilitating molecules" or "CliMs". The present invention further provides genetic constructs for use in genetically modifying eukaryotic or prokaryotic cells and more particularly eukaryotic tissue so as to alter their visual characteristics or capacity for exhibiting same to a human eye in the absence of excitation by an extraneous non-white light or particle emission. The present invention, therefore, extends to eukaryotic or prokaryotic cells and more particularly eukaryotic tissue, which are genetically modified to produce CFMs and which thereby exhibit altered visual characteristics in the absence of excitation by an extraneous non-white light or particle emission. In one particular embodiment, the CFMs are used to alter the visual characteristics of plants and even more particularly flower color. In another embodiment, the present invention provides gels or coatings or similar biomaterials in the form of a biomatrix comprising the CliMs such as for use as a UV sink, in a sun screen, in cosmetics, as an expression marker or other reporter molecule or for use as a photon trap to increase light intensity.



-1-

### CELL VISUAL CHARACTERISTIC-MODIFYING SEQUENCES

#### FIELD OF THE INVENTION

The present invention relates generally to peptides, polypeptides or proteins having one or more amino acids or one or more amino acid sequences which exhibit color-facilitating properties, either on their own or following interaction with one or more other amino acids and to nucleic acid molecules encoding same. Such peptides, polypeptides and proteins are referred to herein as "color-facilitating molecules" or "CFMs". The present invention further provides genetic constructs for use in genetically modifying eukaryotic or prokaryotic cells and more particularly eukaryotic tissue so as to alter their visual characteristics or capacity for exhibiting same to a human eye in the absence of excitation by an extraneous non-white light or particle emission. The present invention, therefore, extends to eukaryotic or prokaryotic cells and more particularly eukaryotic tissue, which are genetically modified to produce CFMs and which thereby exhibit altered visual characteristics in the absence of excitation by an extraneous non-white light or particle emission. In one particular embodiment, the CFMs are used to alter the visual characteristics of plants and even more particularly flower color. In another embodiment, the present invention provides gels or coatings or similar biomaterials in the form of a biomatrix comprising the CFMs such as for use as a UV sink, in a sun screen, in cosmetics, as an expression marker or other reporter molecule or for use as a photon trap to increase light intensity.

#### **BACKGROUND OF THE INVENTION**

25

15

Reference to any prior art in this specification is not, and should not be taken as, an acknowledgment or any form of suggestion that this prior art forms part of the common general knowledge in Australia or any other country.

30 All-protein chromophores (pigments) have been isolated from the phylum Cnidaria (also known as Coelenterata). This phylum contains four classes: Scyphozoa, Cubozoa,

- 2 -

Anthozoa and Hydrozoa. The first all-protein chromophore to be isolated, Green Fluorescent Protein (GFP), was cloned and sequenced from cDNA of the Hydrozoan Aequorea victoria, commonly called jellyfish.

Similar all-protein chromophores have been isolated from Anthozoans. Matz et al. (Nature Biotechnol. 17: 969-973, 1999), used degenerative primers based on Aequorea victoria GFP nucleotide sequence to PCR amplify cDNA isolated from four of the five orders of Anthozoa: Stolonifera, Actiniaria, Zoanthidea, and Corallimorpharia. Lukyanov et al. (Journal of Biological Chemistry 275: 25879-25882, 2000) used the same methodology to isolate a non-fluorescent all-protein chromophore from Actiniaria. However, the methodology used was unable to isolate all-protein chromophores from the fifth order, Scleractinia.

The Scleractinia are corals that form architecture for coral reefs. They are otherwise known as "true" or "reef-building" corals. International Patent Publication No. WO 00/46233 and Dove et al. (Coral Reefs 19: 197-204, 2000) both relate to isolation of an all-protein chromophore derived from Scleractinia pigment protein from coral tissue (PPCT).

All-protein chromophores isolated to date display a range of spectral properties which effect apparent color in specific environments. Color may be determined by absorption and/or fluorescence properties of the molecules as well as qualities of incident light. Spectral properties include absorption, excitation and emission energies, molar extinction coefficients, quantum yields and maturation parameters. In many cases, a simple amino acid substitution can have a dramatic effect on the polypeptide spectral parameters (e.g. Tsien, Ann. Rev. Biochem. 67: 509, 1998; Lukyanov et al., 2000, supra). However, useful modifications of a particular molecule are limited, as directed and random mutagenesis of specific all-protein chromophores has failed to produce desired spectral features (Tsien, 1998, supra). The result is that all-protein chromophores isolated from different sources are finding specific application niches.

. . .

25

15

20

. .

- 3 -

One all-protein chromophores, primarily used as molecular marker, is GFP. This protein, when excited with either UV or blue light (maximally at 396 nm or 475 nm) emits green fluorescence (maximally at 500 nm) [Heim et al., Proc. Natl. Acad. Sci. USA 91: 12501-12504, 1994]. GFP mutants that are altered in their maximal excitation and emission characteristics have been generated by random mutagenesis (Crameri et al., Nature Biotechnology 14: 315-319, 1996). Other GFP mutants have been generated that have increased solubility and fluorescence (Davis and Vierstra, Soluble derivatives of green fluorescent protein (GFP) for use in Arabidopsis thaliana. Weeds of the World, The International Electronic Arabidopsis Newsletter ISSN 1358-6912, (Ed. Mary Anderson) vol 3ii, 1996). The fluorescence of GFP and its mutants has been exploited for noninvasive analysis and monitoring of biological samples in plants and other organisms for research purposes (Haseloff et al., Proc. Natl. Acad. Sci USA 94: 2122-2127, 1997; Hu and Cheng, FEBS Letters 369: 331-334, 1995; Wang and Hazelrigg, Nature 369: 400-403, 1994). The use of these fluorescent GFPs, mutants and homologs as fluorescent marker pigments visible upon excitation by light of specific wavelengths is well documented (e.g. U.S. Patent Nos. 6,027,881 and 5,958,713; Japanese Patent No. 11266883; International Patent Publication No. WO97/11094; U.S. Patent No. 5,625,048; International Patent Application No. PCT/US99/29472 and International Patent Publication PCT/AU00/00056).

20

10

In contrast to other fluorescent proteins, the fluorescence of GFP is due to amino acid interaction within the molecule, generally after folding. A contiguous fluorophore-defining amino acid sequence of Ser-Tyr-Gly is modified upon folding to produce an extended aromatic system which imparts the characteristic green fluorescence to the mature protein (Cody et al., Biochemistry 32: 1212-1218, 1993; Ormö et al., Science 273: 1392-1395, 1996; Yang et al., Nature Biotechnol. 14: 1246-1251, 1996). As stated above, GFP like molecules have been identified for nonbioluminscent Anthozoa species (Matz et al., 1999, supra) which provides evidence that GFP-like proteins are not necessarily components of bioluminescent systems but may just determine fluorescent coloration in animals (Lukyanov et al., 2000, supra). Other weakly fluorescent GFP homologs have been identified from Acropora formosa and Acropora digitifera (Dove et al., Biol. Bull. 189:

-4-

288-297, 1995; Hoegh-Guldberg and Dove, 2000, *supra*; Salih et al., *Nature 408*: 850-853, 2000).

All-protein chromophores are now finding application as molecular markers for monitoring polypeptide expression and localization in the fields of biochemistry, molecular and cell biology.

The present invention now describes novel all-protein chromophores (or CFMs) as well as novel and useful applications of same.

10

15

For example, the flower industry strives to develop new and different varieties of flowering plants, in particular through the manipulation of flower color. While classical breeding techniques have been used with some success to produce a wide range of colors for most of the commercial varieties of flowers, this approach has been limited by the constraints of a particular species' gene pool. For this reason, it is rare for a single species to have a full spectrum of colored varieties. The development of blue varieties of major cut flower species such as rose, chrysanthemum, tulip, lily, carnation and gerbera, for example, has proved difficult and would offer a significant opportunity in both the cut flower and ornamental markets.

20

25

Flower color is predominantly due to three types of pigment: flavonoids, carotenoids and betalains. Of the three, the flavonoids are the most common and contribute to a range of colors from yellow to red to blue. The flavonoid molecules which make the major contribution to flower color are the anthocyanins which are glycosylated derivatives of cyanidin, delphinidin, petunidin, peonidin, malvidin and pelargonidin and are localized in the vacuole. Carotenoids are natural pigments that confer yellow, orange and red colors to flowers and fruit. In plants, these pigments are localized in chromoplasts in flowers, leaves, fruit and roots.

30 Novel colors in ornamental plant and flowering plant species may be generated by modifying the anthocyanin pathway to produce novel anthocyanins and aurones (Davies et

- 5 -

al., Plant Journal 13: 259-266, 1998) and to alter ratios of anthocyanins to co-pigments (Holton et al., Plant Journal 4: 1003-1010, 1993). Alternatively, the carotenoid biosynthetic pathway can be modified to produce novel flower colors (Mann et al., Nature Biotech. 18: 888-892, 2000). The levels of anthocyanin production can also be increased by the expression of heterologous anthocyanin pathway gene regulatory factors (e.g. see Borevitz et al., Plant Cell 12: 2383-2393, 2000).

These approaches have been used with some, albeit limited, success and alternative novel approaches are constantly being sought.

10

15

20

In work leading up to the present invention, the inventors sought, *inter alia*, to identify novel color-facilitating molecules (CFMs) and to use same to modify the visual characteristics of eukaryotic or prokaryotic organisms by introducing into eukaryotic or prokaryotic cells, genetic material encoding CFMs which impart a color visible to a human eye in the absence of excitation by extraneous non-white light or particle emission. In a preferred embodiment, the CFMs are proteins such as GFPs or their relatives, such as non-fluorescent GFP-homologs. The use of CFMs to modulate the color of plants or plant parts such as flowers and seeds, represents a new approach to developing plant varieties having altered color characteristics. Other uses contemplated herein for the CFMs include their use as expression markers or as general reporter molecules, as a photon trap, UV sink and in sun screen or cosmetic or may be embedded in a gel matrix and be used to convert less visible light to wavelengths which are more visible. All such compositions are encompassed by the term "biomatrix".

-6-

#### SUMMARY OF THE INVENTION

Throughout this specification, unless the context requires otherwise, the word "comprise", or variations such as "comprises" or "comprising", will be understood to imply the inclusion of a stated element or integer or group of elements or integers but not the exclusion of any other element or integer or group of elements or integers.

Nucleotide and amino acid sequences are referred to by a sequence identifier number (SEQ ID NO:). The SEQ ID NOs: correspond numerically to the sequence identifiers <400>1, <400>2, etc. A sequence listing is provided after the claims.

The present invention provides peptides, polypeptides and proteins having one or more amino acid sequences which exhibit color-facilitating properties, either on their own or following interaction with one or more amino acids as well as nucleic acid molecules encoding same. Preferably, the peptides, polypeptides and proteins or their nucleic acid molecules are derived from one or more Anemonia majano, Anemonia sulcata, Clavularia sp, Zoanthus sp, Discosoma sp (e.g. Discosoma striata), Aequorea sp (e.g. Aequorea victoria), Anthozoa sp, Cassiopea sp, (e.g. Cassiopea xamachana), Millepora sp, Acropora sp (e.g. Acropora aspera and Acropora nobilis), Montipora sp, Porites murrayensis, Pocillopora damicormis, Pavona descussaca, Acanthastrea sp, Platygyra sp or Caulastrea sp. These peptides, polypeptides and proteins are referred to as "color-facilitating molecules" (CFMs) and may be in isolated form, be produced within or on a cell or may form part of a biomatrix.

Accordingly, in one aspect of the present invention, there is provided an isolated nucleic acid molecule comprising a nucleotide sequence encoding a color-facilitating molecule (CFM) which, in a cell, alone or together with one or more other molecules imparts an altered visual characteristic to said cell when visualized by a human eye in the absence of excitation by extraneous non-white light or particle emission.

-7-

The present invention also provides an isolated CFM comprising a polypeptide which, in a cell, alone or together with one or more other molecules imparts an altered visual characteristic to said cell when visualized by a human eye in the absence of excitation by extraneous non-white light or particle emission.

The preferred CFM comprises the amino-terminal end of the polypeptide set forth in SEQ ID NOs: 5, 6, 7, 8 or 9.

5

15

20

Particularly preferred CFMs comprise amino acid sequences selected from SEQ ID NOs:10, 11, 12, 13, 14, 15, 16, 17 or 18.

Even more preferably, the CFM is encoded by a nucleotide sequence set forth in any one of SEQ ID NOs:19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, 131, 133, 135, 137, 139, 141, 143, 145, 147, 149, 151, 153, 155, 157, 159, 161, 163, 165, 167, 169, 171, 173, 175, 177, 179, 189, 191, 193, 195, 197, 199 and 201 or a nucleotide sequence capable of hybridizing to one of the above sequences or a complementary form thereof under low stringency conditions or a nucleotide sequence having at least about 60% similarity to any one of the above sequences.

Amino acid sequences corresponding to the above nucleotide sequences correpond to SEQ ID NOs:20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178, 180, 190, 192, 194, 196, 198, 200 and 202 as well as an amino acid sequence having at least about 60% similarity to any one of the above sequences.

30 The CFM may be in isolated form or part of a biomatrix wherein the biomatrix includes a cell, solid support, gel or bioinstrument. The CFMs are particularly useful in generating

-8-

eukaryotic or prokaryotic cells exhibiting altered visual characteristics as well as biomatrices in the form of sun screen, UV traps, photon traps and illuminescent intensifiers.

In a particularly preferred embodiment, the present invention provides transgenic plants and parts thereof including flowers, roots, leaves, stems, fruit and fibers exhibiting an altered visual characteristic.

-9-

#### BRIEF DESCRIPTION OF THE FIGURES

Figure 1 shows a representation of multiple alignment of encoded amino acid sequences having SEQ ID NOs:20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68 70, 72, 74, 76, 78, 80, 82, 84 and 86, representing polypeptides comprising an N-terminal SVIAK (SEQ ID NO:5) sequence.

Figure 2 shows corresponding nucleotide sequence alignments of nucleic acid molecules, having SEQ ID NOs:19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83 and 85, encoding the polypeptides shown in Figure 1.

Figure 3 shows a representation of multiple alignment of encoded amino acid sequences having SEQ ID NOs:88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166 and 168, for polypeptides comprising an N-terminal (M)SVIAT (SEQ ID NO:6), SGIAT (SEQ ID NO:7), SVIVT (SEQ ID NO:8) and SVSAT (SEQ ID NO:9) sequences.

Figure 4 shows corresponding nucleotide sequence alignments of nucleic acid molecules, having SEQ ID NOs:87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, 131, 133, 135, 137, 139, 141, 143, 145, 147, 149, 151, 153, 155, 157, 159, 161, 163, 165 and 167, encoding the polypeptides shown in Figures 3A-3D.

25

Figure 5 shows a representation of an alignment of amino acid sequences having SEQ ID NOs:170, 172, 174, 176, 178 and 180, for polypeptides comprising an N-terminal SVIAK sequence (SEQ ID NO:5) and a stop codon corresponding to amino acid residue 14.

PCT/GB02/00928 WO 02/070703

-10-

Figure 6 shows corresponding nucleotide sequence alignments for nucleic acid molecules, having SEQ ID NOs:169, 171, 173, 175, 177 and 179, encoding the polypeptides shown in Figure 5.

Figure 7 is a nucleotide sequence alignment of SEQ ID NO:19 and SEQ ID NO:169, being nucleic acid sequences encoding polypeptides without and with a stop codon corresponding to amino acid residue 14, respectively.

Figure 8 shows a representation of multiple alignment of amino acid sequences for polypeptides comprising an N-terminal SVIAK sequence (SEQ ID NO:5), including SEQ ID NOs:20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68 70, 72, 74, 76, 78, 80, 82, 84 and 86, as well as sequences Aapat-1 (SEQ ID NO:181) and Aapat-2 (SEQ ID NO:182) which are disclosed in International Patent Publication No. WO 00/46233.

15

10

7

 $\mathcal{L}$ 

Figure 9 shows amino acid sequence alignments of pigment polypeptides from coral tissue, grouped according to their N-terminal 5-amino acid sequence. The name and SEQ ID NO for each peptide is indicated, as well as the "Type" to which each has been assigned based on the identity of the 29 amino acids which are located within 5 Angstroms of the "QYG" fluorophore. These 29 individual, non-contiguous amino acid residues are also indicated, as are the individual non-contiguous variable amino acids residues throughout the polypeptides shown.

Figure 10 is a diagrammatic representation of a generic bacterial expression vector based on pQE-30 (Qiagen), into which is inserted an ~0.7kb cDNA; depending on the source of the cDNA clone, each plasmid is designated as follows: pCGP2915 - A10 clone from Acropora sp.; pCGP2916 - All clone from Acropora sp.; pCGP2917 - Al2 clone from Acropora sp.; pCGP2918 - A8 clone from Acropora sp. (SEQ ID NO:189); pCGP2920 -D10 clone from Discosoma sp. (SEQ ID NO:191); pCGP2922 - T3 clone from Tubastrea sp. (SEQ ID NO:195); pCGP2924 - S3 clone from Sinularia sp. (SEQ ID NO:193); pCGP2919 - D1 clone from Discosoma sp. (SEQ ID NO:197); pCGP2921 - T1 clone from

Tubastrea sp. (SEQ ID NO:201); pCGP2923 - S1 clone from Sinularia sp. (SEQ ID NO:199). Abbreviations are as follows: bla =  $\beta$ -lactamase gene; ColE1ori = plasmid origin of replication. The locations of restriction endonuclease recognition sites for PstI, HindIII and BamHI are also marked. Refer to Example 3 for further details.

5

10

Figure 11 is a graphical representation of examples of absorption scans of five "Type 1" (refer to text in Example 2 and Tables 6 and 7 for further detail) colored proteins showing extinction coefficients ( $\varepsilon_{\lambda max}$ ) based on the method of Whitaker and Granum, 1980 (*Anal. Biochem.* 109:156-159) for calculating protein concentration. x-axis = relative absorption; y-axis = wavelength (nm); (a) Rtms5.pep (SEQ ID NO:166), where  $\varepsilon_{592} = 111,000 \text{ M}^{-1} \text{ cm}^{-1}$ ; (b) LGasv-C.pep (SEQ ID NO:44) where  $\varepsilon_{591} = 53,000 \text{ M}^{-1} \text{ cm}^{-1}$ ; (c) Ce61-7sv.pep (SEQ ID NO:38) where  $\varepsilon_{591.5} = 104,000 \text{ M}^{-1} \text{ cm}^{-1}$ ; (d) PPd57-2ms.pep (SEQ ID NO:140) where  $\varepsilon_{593} = 67,000 \text{ M}^{-1} \text{ cm}^{-1}$ ; (e) Mims-C.pep (SEQ ID NO:126) where  $\varepsilon_{589} = 48,000 \text{ M}^{-1} \text{ cm}^{-1}$ .

- 15 Figure 12 a graphical representation of examples of absorption scans of three "Type 2"
  (A) and two "Type 12" (B) (refer to text in Example 2 and Tables 6 and 7 for further detail) colored proteins, showing extinction coefficients (ε λmax) based on the method of Whitaker and Granum (Anal. Biochem. 109: 156-159, 1980) for calculating protein concentration. x-axis = relative absorption; y-axis = wavelength (nm); (A) (a) PMms-20 B.pep (SEQ ID NO:130) where ε<sub>579.5</sub> = 39,000 M<sup>-1</sup> cm<sup>-1</sup>; (b) LGAsv-D.pep (SEQ ID NO:46) where ε<sub>579.5</sub> = 72,400 M<sup>-1</sup> cm<sup>-1</sup>; (c) rtsv-2.pep (SEQ ID NO:84) where ε<sub>579.5</sub> = 75,000 M<sup>-1</sup> cm<sup>-1</sup>; (B) (d) Misv-F.pep (SEQ ID NO:54) where ε<sub>579</sub> = 111,000 M<sup>-1</sup> cm<sup>-1</sup>; (e) Acasv-C.pep (SEQ ID NO:78) where ε<sub>579.5</sub> = 32,300 M<sup>-1</sup> cm<sup>-1</sup>.
- Figure 13 a graphical representation of examples of absorption scans of two "Type 6" (refer to text in Example 2 and Tables 6 and 7 for further detail) colored proteins, showing extinction coefficients (ε λmax) based on the method of Whitaker and Granum (Anal. Biochem. 109: 156-159, 1980) for calculating protein concentration. x-axis = relative absorption; y-axis = wavelength (nm); (a) LGAms-5.pep (SEQ ID NO:116) where ε<sub>583.5</sub> = 71,000 M<sup>-1</sup> cm<sup>-1</sup>; (b) Rtms-1.pep (SEQ ID NO:162) where ε<sub>584</sub> = 44,000 M<sup>-1</sup> cm<sup>-1</sup>.

10

20

Figure 14 a graphical representation of (A) Absorption spectra and (B) Chromatogram of gel filtrated protein elution, both showing 95% confidence intervals for N = 5, for raw phosphate buffer extract of two colour morphs of *Acropora aspera* (dark blue pigmented morph; cream morph). In (A), the estimation of blue-purple pocilloporin concentration per surface area of coral tissue is based on an extinction coefficient range of 50,000 – 100,000 M<sup>-1</sup>cm<sup>-1</sup>. In (B), the chromatogram of gel filtrated protein elution is determined from 235 nm chromatograms and 280 nm chromatograms, applying the equation: 235nm -280 nm)/ 2.51 (Whitaker and Granum, 1980, *supra*). The total area under the graph represents the total soluble protein. Blue-purple pocilloporin concentration is based on the difference between areas under the blue and cream graph in the range of pocilloporin elution (24 - 26.5 min).

Figure 15 is a representation of multiple alignment of encoded amino acid sequences from T1 (SEQ ID NO:202), D1 (SEQ ID NO:198), S1 (SEQ ID NO:200), T3 (SEQ ID NO:196), D10 (SEQ ID NO:192), S3 (SEQ ID NO:194) and A8 (SEQ ID NO:190).

Figure 16 is a representation of multiple alignment of encoded amino acid sequences from SVIAK (SEQ ID NO:5)-containing peptides T1 (SEQ ID NO:202), D1 (SEQ ID NO:198), S1 (SEQ ID NO:200), T3 (SEQ ID NO:196), D10 (SEQ ID NO:192), S3 (SEQ ID NO:194) and A8 (SEQ ID NO:190), together with the SVIAK (SEQ ID NO:5)-containing peptides shown in Figure 1, having SEQ ID NOs:20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68 70, 72, 74, 76, 78, 80, 82, 84 and 86.

Figure 17 is a diagrammatic representation of the yeast expression plasmid pCGP3269. The T1 cDNA (SEQ ID NO:201) cloned in a sense orientation behind the yeast glyceraldehyde 3-phosphate dehydrogenase promoter (PGAP) in the expression vector pYE22m. Abbreviations are as follows: TRP1 = Trp1 gene, TGAP = terminator sequence from the yeast glyceraldehyde 3-phosphate dehydrogenase gene, IR1 = inverted repeat of 2 μm plasmid, pBR322 = origin of replication from E. coli. A selection of restriction enonuclase recognition sites are also marked. Refer to Example 7 for further details.

- 13 -

Figure 18 is a diagrammatic representation of the yeast expression plasmid pCGP3270. The A8 cDNA (SEQ ID NO:189) cloned in a sense orientation behind the yeast glyceraldehyde 3-phosphate dehydrogenase promoter (PGAP) in the expression vector pYE22m. Abbreviations are as follows: TRP1 = Trp1 gene, TGAP = terminator sequence from the yeast glyceraldehyde 3-phosphate dehydrogenase gene, IR1 = inverted repeat of 2 μm plasmid, pBR322 = origin of replication from E. coli. A selection of restriction enonuclase recognition sites are also marked. Refer to Example 7 for further details.

Figure 19 is a diagrammatic representation of a plasmid, designated pCGP2756, which comprises a multiple cloning site from pNEB193 (New England Biolabs) between the CaMV (Cauliflower Mosaic Virus) 35S promoter and CaMV 35S terminator sequences. Abbreviations are as follows: Amp = ampicillin resistance gene; p35S = a promoter region from the CaMV 35S gene; t35S = a terminator fragment from the CaMV 35S gene. A selection of restriction endonuclease recognition sites are also marked. Refer to Example 9 for further details.

Figure 20 is a diagrammatic representation of the binary plasmid pCGP2757, which comprises the CaMV35S expression cassette of pCGP2756 (Figure 19) and a SuRB selectable marker gene. Abbreviations are as follows: TetR = the tetracycline resistance gene; LB = left border; RB = right border; SuRB = the coding region and terminator sequence from the acetolactate synthase gene from tobacco; p35S = a promoter region from the cauliflower mosaic virus (CaMV) 35S gene; t35S = a terminator fragment from the CaMV 35S gene; pVS1 = a broad host range origin of replication from a plasmid from Pseuodomonas aeruginosa; pACYC ori = modified replicon from pACYC184 from E. coli. Selected restriction endonuclease recognition sites are also marked. Refer to Example 9 for further details.

20

25

Figure 21 is a diagrammatic representation of the binary plasmid pCGP2765, which comprises the A8 cDNA from *Acropora* sp. (SEQ ID NO:189) cloned into the binary vector pCGP2757 (Figure 20). Abbreviations are as follows: TetR = the tetracycline

- 14 -

resistance gene; LB = left border; RB = right border; SuRB = the coding region and terminator sequence from the acetolactate synthase gene from tobacco; p35S = a promoter region from the cauliflower mosaic virus (CaMV) 35S gene; t35S = a terminator fragment from the CaMV 35S gene; pVS1 = a broad host range origin of replication from a plasmid from Pseuodomonas aeruginosa; pACYC ori = modified replicon from pACYC184 from E. coli; A8 = cDNA from Acropora sp. (SEQ ID NO:189). Selected restriction endonuclease recognition sites are also marked. Refer to Example 9 for further details.

Figure 22 is a diagrammatic representation of the binary plasmid pCGP2769, which comprises the D1 cDNA from *Discosoma* sp. (SEQ ID NO:197) cloned into the binary vector pCGP2757 (Figure 20). Abbreviations are as follows: TetR = the tetracycline resistance gene; LB = left border; RB = right border; SuRB = the coding region and terminator sequence from the acetolactate synthase gene from tobacco; p35S = a promoter region from the cauliflower mosaic virus (CaMV) 35S gene; t35S = a terminator fragment from the CaMV 35S gene; pVS1 = a broad host range origin of replication from a plasmid from Pseuodomonas aeruginosa; pACYC ori = modified replicon from pACYC184 from E. coli; D1 = cDNA from Discosoma sp. (SEQ ID NO:197). Selected restriction endonuclease recognition sites are also marked. Refer to Example 9 for further details.

Figure 23 is a diagrammatic representation of the binary plasmid pCGP2770, which comprises the S1 cDNA from Sinularia sp. (SEQ ID NO:199) cloned into the binary vector pCGP2757 (Figure 20). Abbreviations are as follows: TetR = the tetracycline resistance gene; LB = left border; RB = right border; SuRB = the coding region and terminator sequence from the acetolactate synthase gene from tobacco; p35S = a promoter region from the cauliflower mosaic virus (CaMV) 35S gene; t35S = a terminator fragment from the CaMV 35S gene; pVS1 = a broad host range origin of replication from a plasmid from Pseuodomonas aeruginosa; pACYC ori = modified replicon from pACYC184 from E. coli; S1 = cDNA from Sinularia sp. (SEQ ID NO:199). Selected restriction endonuclease recognition sites are also marked. Refer to Example 9 for further details.

5

10

15

- 15 -

Figure 24 is a diagrammatic representation of the binary plasmid pCGP2772, which comprises the T1 cDNA from Tubastrea sp. (SEQ ID NO:201) cloned into the binary vector pCGP2757 (Figure 20). Abbreviations are as follows: TetR = the tetracycline resistance gene; LB = left border, RB = right border; SuRB = the coding region and 5 terminator sequence from the acetolactate synthase gene from tobacco; p35S = a promoter region from the cauliflower mosaic virus (CaMV) 35S gene; t35S = a terminator fragment from the CaMV 35S gene; pVS1 = a broad host range origin of replication from a plasmid from Pseuodomonas aeruginosa; pACYC ori = modified replicon from pACYC184 from E. coli; T1 = cDNA from Tubastrea sp. (SEQ ID NO:201). Selected restriction endonuclease recognition sites are also marked. Refer to Example 9 for further details.

Figure 25 is a diagrammatic representation of the plasmid pCGP1116, which comprises a promoter fragment from a chalcone synthase (CHS) gene from Rosa hybrida cv. Kardinal. Abbreviations are as follows: Rose CHS = Rose chalcone synthase promoter fragment; ori = origin of replication; Amp = ampicillin resistance gene; Several restriction endonuclease recognition sites are also marked. Refer to Example 10 for further details.

10

15

20

Figure 26 is a diagrammatic representation of the binary plasmid pCGP3255. The CaMV35S promoter of the 35S expression cassette of pCGP2757 (Figure 20) has been replaced with the rose chalcone synthase promoter fragment from pCGP1116 (Figure 25) Abbreviations are as follows: rCHS = rose chalcone synthase promoter fragment; TetR = the tetracycline resistance gene; LB = left border; RB = right border; SuRB = the coding region and terminator sequence from the acetolactate synthase gene from tobacco; p35S = a promoter region from the cauliflower mosaic virus (CaMV) 35S gene; t35S = a terminator fragment from the CaMV 35S gene; pVS1 = a broad host range origin of replication from a plasmid from Pseuodomonas aeruginosa; pACYC ori = modified replicon from pACYC184 from E. coli. Refer to Example 10 for further details.

Figure 27 is a diagrammatic representation of the bianry plasmid pCGP2782. The T1 cDNA from Tubastrea sp. (SEQ ID NO:201) was cloned into binary vector pCGP3255 30

- 16 -

(Figure 26) behind the rose chalcone synthase promoter fragment. Abbreviations are as follows: rCHS = rose chalcone synthase promoter fragment; TetR = the tetracycline resistance gene; LB = left border; RB = right border; SuRB = the coding region and terminator sequence from the acetolactate synthase gene from tobacco; p35S = a promoter region from the cauliflower mosaic virus (CaMV) 35S gene; t35S = a terminator fragment from the CaMV 35S gene; pVS1 = a broad host range origin of replication from a plasmid from Pseuodomonas aeruginosa; pACYC ori = modified replicon from pACYC184 from E. coli; T1 = cDNA from Tubastrea sp. (SEQ ID NO:201). A selection of restriction endomuclease recognition sites is also marked. Refer to Example 10 for further details.

10

15

20

Figure 28 is a diagrammatic representation of the binary plasmid pCGP2773. The D1 cDNA from *Discosoma* sp. (SEQ ID NO:197) was cloned into binary vector pCGP3255 (Figure 26), behind the rose chalcone synthase promoter fragment. Abbreviations are as follows: rCHS = rose chalcone synthase promoter fragment; TetR = the tetracycline resistance gene; LB = left border; RB = right border; SuRB = the coding region and terminator sequence from the acetolactate synthase gene from tobacco; p35S = a promoter region from the cauliflower mosaic virus (CaMV) 35S gene; t35S = a terminator fragment from the CaMV 35S gene; pVS1 = a broad host range origin of replication from a plasmid from Pseuodomonas aeruginosa; pACYC ori = modified replicon from pACYC184 from E. coli; D1 = cDNA from Discosoma sp. (SEQ ID NO:197). A selection of restriction endonuclease recognition sites is also marked. Refer to Example 10 for further details.

Figure 29 is a diagrammatic representation of the binary plasmid pCGP2774. The S1 cDNA from Sinularia sp. (SEQ ID NO:199) was cloned into binary vector pCGP3255 (Figure 26), behind the rose chalcone synthase promoter fragment. Abbreviations are as follows: rCHS = rose chalcone synthase promoter fragment; TetR = the tetracycline resistance gene; LB = left border; RB = right border; SuRB = the coding region and terminator sequence from the acetolactate synthase gene from tobacco; p35S = a promoter region from the cauliflower mosaic virus (CaMV) 35S gene; t35S = a terminator fragment from the CaMV 35S gene; pVS1 = a broad host range origin of replication from a plasmid

- 17 -

from *Pseuodomonas aeruginosa*; pACYC ori = modified replicon from pACYC184 from *E. coli*; S1 = cDNA from *Sinularia* sp. (SEQ ID NO:199). A selection of restriction endonuclease recognition sites is also marked. Refer to Example 10 for further details.

Figure 30 is a diagrammatic representation of the binary plasmid pCGP2780, which is plasmid pCGP2757 (Figure 20) from which has been removed a ~290 base-pair Sall fragment to allow the creation of a unique BamHI restriction endonuclease site. Abbreviations are as follows: TetR = the tetracycline resistance gene; LB = left border; RB = right border; SuRB = the coding region and terminator sequence from the acetolactate synthase gene from tobacco; p35S = a promoter region from the cauliflower mosaic virus (CaMV) 35S gene; t35S = a terminator fragment from the CaMV 35S gene; pVS1 = a broad host range origin of replication from a plasmid from Pseuodomonas aeruginosa; pACYC ori = modified replicon from pACYC184 from E. coli A selection of restriction endonuclease recognition sites is also marked. Refer to Example 11 for further details.

15

20

Figure 31 is a diagrammatic representation of the binary plasmid pCGP2784, which is comprised of the ~0.2 kb chloroplast transit-peptide from the small subunit of ribulose bisphosphate carboxylase gene (RBCase) from *Nicotiana sylvestris*, cloned into the multiple cloning site of pCGP2780 of Figure 30. Abbreviations are as follows: TetR = the tetracycline resistance gene; LB = left border; RB = right border; SuRB = the coding region and terminator sequence from the acetolactate synthase gene from tobacco; p35S = a promoter region from the cauliflower mosaic virus (CaMV) 35S gene; t35S = a terminator fragment from the CaMV 35S gene; pVS1 = a broad host range origin of replication from a plasmid from *Pseuodomonas aeruginosa*; pACYC ori = modified replicon from pACYC184 from *E. coli*; TSSU = chloroplast transit-peptide from the small subunit of RBCase of *Nicotiana sylvestris*. Selected restriction endonuclease recognition sites are also marked. Refer to Example 11 for further details.

Figure 32 is a diagrammatic representation of the binary plasmid pCGP2781, which is plasmid pCGP2772 (Figure 24) from which has been removed a ~290 base-pair Sall

- 18 -

fragment to allow the creation of a unique BamHI restriction endonuclease site. Abbreviations are as follows: TetR = the tetracycline resistance gene; LB = left border; RB = right border; SuRB = the coding region and terminator sequence from the acetolactate synthase gene from tobacco; p35S = a promoter region from the cauliflower mosaic virus (CaMV) 35S gene; t35S = a terminator fragment from the CaMV 35S gene; pVS1 = a broad host range origin of replication from a plasmid from Pseudomonas aeruginosa; pACYC ori = modified replicon from pACYC184 from E. coli. T1 = T1 cDNA from Tubastrea sp. (SEQ ID NO:201). Selected restriction endonuclease recognition sites are also marked. Refer to Example 11 for further details.

10

15

20

Figure 33 is a diagrammatic representation of the binary plasmid pCGP2785, which is comprised of the ~0.2 kb chloroplast transit peptide from the small subunit of ribulose biphosphate carboxylase (RBCase) from *Nicotiana sylvestris* inserted into the CaMV 35S expression cassette of binary vector pCGP2781 (Figure 32), upstream of the T1 cDNA. Abbreviations are as follows: TetR = the tetracycline resistance gene; LB = left border; RB = right border; *SuRB* = the coding region and terminator sequence from the acetolactate synthase gene from tobacco; p35S = a promoter region from the cauliflower mosaic virus (CaMV) 35S gene; t35S = a terminator fragment from the CaMV 35S gene; pVS1 = a broad host range origin of replication from a plasmid from *Pseuodomonas aeruginosa*; pACYC ori = modified replicon from pACYC184 from *E. coli*. T1 = T1 cDNA from *Tubastrea* sp. (SEQ ID NO:201); TSSU = chloroplast transit peptide from the small subunit of RBCase from *Nicotiana sylvestris*. Selected restriction endonuclease recognition sites are also marked. Refer to Example 11 for further details.

Figure 34 is a diagrammatic representation of the binary plasmid pCGP2787 which is comprised of the ~0.2 kb chloroplast transit peptide from the small subunit of ribulose biphosphate carboxylase (RBCase) from *Nicotiana sylvestris* inserted into the Rose CHS expression cassette of binary vector pCGP2782 (Figure 27), upstream of the T1 cDNA. Abbreviations are as follows: TetR = the tetracycline resistance gene; LB = left border; RB = right border; SuRB = the coding region and terminator sequence from the acetolactate

- 19 -

į,

synthase gene from tobacco; p35S = a promoter region from the cauliflower mosaic virus (CaMV) 35S gene; t35S = a terminator fragment from the CaMV 35S gene; rCHS = rose chalcone synthase promoter fragment; pVS1 = a broad host range origin of replication from a plasmid from *Pseuodomonas aeruginosa*; pACYC ori = modified replicon from pACYC184 from *E. coli*. T1 = T1 cDNA from *Tubastrea* sp. (SEQ ID NO:201); TSSU = chloroplast transit peptide from the small subunit of RBCase from *Nicotiana sylvestris*. Selected restriction endonuclease recognition sites are also marked. Refer to Example 11 for further details.

of the basic chitinase N-terminal endoplasmic reticulum (ER) transit peptide signal sequence from Arabidopsis thaliana inserted into the CaMV 35S expression cassette of binary vector pCGP2780 (Figure 30), downstream of the CaMV 35S promoter. Abbreviations are as follows: TetR = the tetracycline resistance gene; LB = left border; RB = right border; SuRB = the coding region and terminator sequence from the acetolactate synthase gene from tobacco; p35S = a promoter region from the cauliflower mosaic virus (CaMV) 35S gene; t35S = a terminator fragment from the CaMV 35S gene; pVS1 = a broad host range origin of replication from a plasmid from Pseuodomonas aeruginosa; pACYC ori = modified replicon from pACYC184 from E. coli; ERT = ER transit peptide signal sequence from Arabidopsis basic chitinase gene. Selected restriction endonuclease recognition sites are also marked. Refer to Example 11 for further details.

Figure 36 is a diagrammatic representation of the binary plasmid pCGP3259. The T1 cDNA from *Tubastrea* sp. (SEQ ID NO:201) with an in-frame HDEL peptide sequence at the 3' end was cloned into the CaMV 35S expression cassette of binary vector pCGP3257 (Figure 35), downstream of the ER transit-peptide signal sequence from *Arabidopsis thaliana*. Abbreviations are as follows: TetR = the tetracycline resistance gene; LB = left border; RB = right border; SuRB = the coding region and terminator sequence from the acetolactate synthase gene from tobacco; p35S = a promoter region from the cauliflower mosaic virus (CaMV) 35S gene; t35S = a terminator fragment from the CaMV 35S gene;

- 20 -

pVS1 = a broad host range origin of replication from a plasmid from *Pseuodomonas* aeruginosa; pACYC ori = modified replicon from pACYC184 from *E. coli*; ERT:T1:HDEL = T1 cDNA clone from *Tubastrea* (SEQ ID NO:201) with an in-frame ER transit peptide sequence from *Arabidopsis* basic chitinase gene at the 5' end and an HDEL ER retention sequence at the 3' end. Selected restriction endonuclease recognition sites are also marked. Refer to Example 11 for further details.

Figure 37 is a diagrammatic representation of the binary plasmid pCGP3262 which is comprised of the basic chitinase N-terminal endoplasmic reticulum (ER) transit peptide signal sequence from Arabidopsis thaliana inserted into the Rose CHS expression cassette of binary vector pCGP3255 (Figure 26), downstream of the Rose CHS promoter. Abbreviations are as follows: TetR = the tetracycline resistance gene; LB = left border; RB = right border; SuRB = the coding region and terminator sequence from the acetolactate synthase gene from tobacco; p35S = a promoter region from the cauliflower mosaic virus (CaMV) 35S gene; t35S = a terminator fragment from the CaMV 35S gene; rCHS rose chalcone synthase promoter fragment; pVS1 = a broad host range origin of replication from a plasmid from Pseuodomonas aeruginosa; pACYC ori = modified replicon from pACYC184 from E. coli; ERT = ER transit peptide signal sequence from Arabidopsis basic chitinase gene. Selected restriction endonuclease recognition sites are also marked. Refer to Example 11 for further details.

10

15

20

Figure 38 is a diagrammatic representation of the binary plasmid pCGP3263. The T1 cDNA from Tubastrea sp. (SEQ ID NO:201) with an in-frame HDEL peptide sequence at the 3' end was cloned into the Rose CHS expression cassette of binary vector pCGP3262 (Figure 37), downstream of the ER transit-peptide signal sequence from Arabidopsis thaliana. Abbreviations are as follows: TetR = the tetracycline resistance gene; LB = left border; RB = right border; SuRB = the coding region and terminator sequence from the acetolactate synthase gene from tobacco; p35S = a promoter region from the cauliflower mosaic virus (CaMV) 35S gene; t35S = a terminator fragment from the CaMV 35S gene; pVS1 = a broad host range origin of replication from a plasmid from Pseuodomonas

- 21 -

aeruginosa; pACYC ori = modified replicon from pACYC184 from E. coli; ERT:T1:HDEL = T1 cDNA clone from Tubastrea (SEQ ID NO:201) with an in-frame ER transit peptide sequence from Arabidopsis basic chitinase gene at the 5' end and an HDEL ER retention sequence at the 3' end; rCHS = Rose chalcone synthase promoter fragment. Selected restriction endonuclease recognition sites are also marked. Refer to Example 11 for further details.

5

10

15

Figure 39 is a diagrammatic representation of the binary plasmid pCGP3258. An in-frame fusion of the T1 coding sequence (SEQ ID NO:201) and the mgfp4 sequence was cloned into the CaMV 35S expression cassette of pCGP3257 (Figure 35). Abbreviations are as follows: TetR = the tetracycline resistance gene; LB = left border; RB = right border; SuRB = the coding region and terminator sequence from the acetolactate synthase gene from tobacco; p35S = a promoter region from the cauliflower mosaic virus (CaMV) 35S gene; t35S = a terminator fragment from the CaMV 35S gene; pVS1 = a broad host range origin of replication from a plasmid from Pseuodomonas aeruginosa; pACYC ori = modified replicon from pACYC184 from E. coli; T1:mgfp4 = T1 cDNA clone from Tubastrea (SEQ ID NO:201) with an in-frame fusion of the mgfp4 coding sequence. Selected restriction endonuclease recognition sites are also marked. Refer to Example 12 for further details.

Figure 40 is a representation of an autoradiograph of an RNA blot probed with <sup>32</sup>P-labelled fragments of (A) a 0.7 kb BamHI/HindIII fragment of the T1 clone contained in pCGP2921 (Figure 10) and (B) 0.8 kb HindIII fragment of SuRB contained in pCGP1651. Each lane contained a 5 to 10 μg sample of total RNA isolated from the leaves and petals of transgenic P. hybrida plants. (C) Ethidium bromide staining of the 18S rRNA is shown as an indication of RNA loading levels. Lane numbers are marked 1 to 12. The numbers above the lane numbers refer to construct pCGP numbers used in the transformation experiments. Refer to Example 15 for further details.

Figure 41 is a representation of an autoradiograph of an RNA blot probed with <sup>32</sup>P-30 labelled fragments of (A) a 0.7 kb *BamHI/HindIII* fragment of the T1 clone contained in

- 22 -

pCGP2921 (Figure 10) and (B) 0.8 kb HindIII fragment of SuRB contained in pCGP1651. Each lane contained a 5 µg sample of total RNA isolated from the leaves of non-transgenic and transgenic A. thaliana plants. (C) Ethidium bromide staining of the 25S rRNA is shown as an indication of RNA loading levels. Lane numbers are marked 1 to 17. The numbers above the lane numbers refer to construct pCGP numbers used in the transformation experiments with the exception of NTG and 35Smgfp4. NTG = non transgenic; 35Smgfp4 = pBIN35Smgfp4. Refer to Example 14 for further details.

Figure 42 is a graphical representation of absorption, excitation and emission spectra for Rtms-5 (SEQ ID NO:166) and its variants. (A) Absorption spectra for Rtms-5 (SEQ ID NO:166); (B) Absorption spectra for variants generated via site directed mutagenesis: Rtms5-H142S and Rtms-5v (SEQ ID NO:216); C Excitation (exc) and emission (em) spectra for Rtms5-H142S and Rtms-5v (SEQ ID NO:216) at wavelengths indicated.

Figure 43 is a graphical representation of examples of excitation and emission spectra for two other colored proteins, showing extinction coefficients (ε λmax) based on the method of Whitaker and Granum (1980, supra) for calculating protein concentration. x-axis = relative absorption; y-axis = wavelength (nm); (A) Aams-4 (SEQ ID NO:90)-H142S, and (B) Rtms-1 (SEQ ID NO:162)-N142S; λmax for each spectrum is shown on the figure.

20

25

Figure 44 is a diagrammatic representation of the binary plasmid pCGP2926. A ~0.1kb AscI/BamHI fragment (containing sequences to a prokaryotic ribosome binding site (RBS), translational initiation consensus sequence (TICS) and an RGSHHHHHHH epitope) generated by ligating the primers TICS-His-FWD (SEQ ID NO:227) and TICS-His-REV (SEQ ID NO:228) was introduced into the binary plasmid pCGP2781 (Figure 32). Abbreviations are as follows: TetR = the tetracycline resistance gene; LB = left border; RB = right border; SuRB = the coding region and terminator sequence from the acetolactate synthase gene from tobacco; p35S = a promoter region from the cauliflower mosaic virus (CaMV) 35S gene; t35S = a terminator fragment from the CaMV 35S gene; pVS1 = a broad host range origin of replication from a plasmid from Pseuodomonas aeruginosa;

pACYC ori = modified replicon from pACYC184 from E. coli. T1 = T1 cDNA from

- 23 -

Tubastrea sp. (SEQ ID NO:201), His = RGSHHHHHHH epitope. Selected restriction endonuclease recognition sites are also marked. Refer to Example 9 for further details.

Figure 45 is diagrammatic representation of the binary plasmid pCGP3261. An ER targeted T1:mGFP4 fusion was cloned into CaMV 35S expression cassette of the binary vector pCGP3257. Abbreviations are as follows: TetR = the tetracycline resistance gene; LB = left border; RB = right border; SuRB = the coding region and terminator sequence from the acetolactate synthase gene from tobacco; p35S = a promoter region from the cauliflower mosaic virus (CaMV) 35S gene; t35S = a terminator fragment from the CaMV 35S gene; pVS1 = a broad host range origin of replication from a plasmid from Pseuodomonas aeruginosa; pACYC ori = modified replicon from pACYC184 from E. coli; ERT:T1:mGFP4:HDEL = T1 cDNA clone from Tubastrea (SEQ ID NO:201):mGFP4 in-frame fusion with an in-frame ER transit peptide sequence from Arabidopsis basic chitinase gene at the 5' end and an HDEL ER retention sequence at the 3' end. Selected restriction endonuclease recognition sites are also marked. Refer to Example 12 for further details.

Figure 46 is diagrammatic representation of the binary plasmid pCGP3260. An ER targeted mGFP4 coding region was cloned into CaMV 35S expression cassette of the binary vector pCGP2780. Abbreviations are as follows: TetR = the tetracycline resistance gene; LB = left border; RB = right border; SuRB = the coding region and terminator sequence from the acetolactate synthase gene from tobacco; p35S = a promoter region from the cauliflower mosaic virus (CaMV) 35S gene; t35S = a terminator fragment from the CaMV 35S gene; pVS1 = a broad host range origin of replication from a plasmid from Pseuodomonas aeruginosa; pACYC ori = modified replicon from pACYC184 from E. coli; ERT:mGFP4:HDEL = mGFP4 coding sequence with an in-frame ER transit peptide sequence from Arabidopsis basic chitinase gene at the 5' end and an HDEL ER retention sequence at the 3' end. Selected restriction endonuclease recognition sites are also marked. Refer to Example 12 for further details.

- 24 -

Figure 47 is a photographic representation of clear nature gel electrophoresis showing separation of fluorescently labeled mitochondrial ATP synthase. 1. b-gfp fusion protein; 2. b-Rtms-5v fusion protein; 3. b-dsRed fusion protein; 4. GFP not fused to another protein.

A summary of sequence identifiers used throughout the subject specification is provided in Table 1.

TABLE 1
SUMMARY OF SEQUENCE IDENTIFIERS

5

SEQ	NAME	DESCRIPTION
ID		
NO.		
1	POC FOR	oligonucleotide
2	POC 220	oligonucleotide
3	MSVIAT FOR	oligonucleotide
4	POC 231	oligonucleotide
5	SVIAK	N-terminal amino acid sequence of a CFM
6	(M)SVIAT	N-terminal amino acid sequence of a CFM
7	SGIAT	N-terminal amino acid sequence of a CFM
8	SVIVT	N-terminal amino acid sequence of a CFM
9	SVSAT	N-terminal amino acid sequence of a CFM
10	SVIATQMTYKVYMSGT	N-terminal amino acid sequence of a CFM
11	SVIATQMTYKVYMSPT	N-terminal amino acid sequence of a CFM
12	SVIATQVTYKVYMSGT	N-terminal amino acid sequence of a CFM
13	SGIATQMTYKVYMSGT	N-terminal amino acid sequence of a CFM
14	SVIVTQMTYKVYMSGT	N-terminal amino acid sequence of a CFM
15	SVSATQMTYKVYMSG	N-terminal amino acid sequence of a CFM
	T	
16	SVIAKQMTYKVNMSG	N-terminal amino acid sequence of a CFM
1	1	
17	SVIAKQMTYKVYMSD	N-terminal amino acid sequence of a CFM
	T	
18	SVIAKQMTYX <sub>1</sub> X <sub>2</sub> YX <sub>3</sub> S	N-terminal amino acid sequence of a CFM
19	GT Aasv-1	CVIAV to colore
20		nucleotide sequence of SVIAK-type clone
21	Aasv-1.pep Aasv-3	translated amino acid sequence of SVIAK CFM
22		nucleotide sequence of SVIAK-type clone
23	Aasv-3.pep Aasv-P	translated amino acid sequence of SVIAK CFM
		nucleotide sequence of SVIAK-type clone
24	Aasv-P.pep	translated amino acid sequence of SVIAK CFM
25	Acasv-A	nucleotide sequence of SVIAK-type clone
26	Acasv-A.pep	translated amino acid sequence of SVIAK CFM
27	Acasv-C	nucleotide sequence of SVIAK-type clone
28	Acasv-C.pep	translated amino acid sequence of SVIAK CFM

SEQ	NAME	DESCRIPTION
ID	NAUVLE	DESCRIPTION
NO.		
29	Acasv-D	nucleotide sequence of SVIAK-type clone
30	Acasy-D.pep	translated amino acid sequence of SVIAK CFM
31	Ce61-3sv	nucleotide sequence of SVIAK-type clone
32	Ce61-3sv.pep	translated amino acid sequence of SVIAK CFM
33	Ce61-4sy	nucleotide sequence of SVIAK-type clone
34	Ce61-4sv.pep	translated amino acid sequence of SVIAK CFM
35	Ce61-5sy	nucleotide sequence of SVIAK-type clone
36	Ce61-5sv.pep	translated amino acid sequence of SVIAK CFM
37	Ce61-7sv	nucleotide sequence of SVIAK-type clone
38	Ce61-7sv.pep	translated amino acid sequence of SVIAK CFM
39	GPd58-2sv	nucleotide sequence of SVIAK-type clone
40	GPd58-2sv.pep	translated amino acid sequence of SVIAK CFM
41	LGAsv-A	nucleotide sequence of SVIAK-type clone
42	LGAsv-A.pep	translated amino acid sequence of SVIAK CFM
43	LGAsv-C	nucleotide sequence of SVIAK-type clone
44	LGAsv-C.pep	translated amino acid sequence of SVIAK CFM
45	LGAsv-D	nucleotide sequence of SVIAK-type clone
46	LGAsv-D.pep	translated amino acid sequence of SVIAK CFM
47	LGAsv-E	nucleotide sequence of SVIAK-type clone
48	LGAsv-E.pep	translated amino acid sequence of SVIAK CFM
49	Misv-A	nucleotide sequence of SVIAK-type clone
50	Misv-A.pep	translated amino acid sequence of SVIAK CFM
51	Misv-B	nucleotide sequence of SVIAK-type clone
52	Misv-B.pep	translated amino acid sequence of SVIAK CFM
53	Misv-F	nucleotide sequence of SVIAK-type clone
54	Misv-F.pep	translated amino acid sequence of SVIAK CFM
55	PM1Asv-rep	nucleotide sequence of SVIAK-type clone
56	PM1Asv-rep.pep	translated amino acid sequence of SVIAK CFM
57	PM1Csv-rep	nucleotide sequence of SVIAK-type clone
58	PM1Csv-rep.pep	translated amino acid sequence of SVIAK CFM
59	PMsv-4	nucleotide sequence of SVIAK-type clone
60	PMsv-4.pep	translated amino acid sequence of SVIAK CFM
61	PMsv-5	nucleotide sequence of SVIAK-type clone
62	PMsv-5.pep	translated amino acid sequence of SVIAK CFM
63	PPsv-1	nucleotide sequence of SVIAK-type clone
64	PPsv-1.pep	translated amino acid sequence of SVIAK CFM
65	PPsv-2	nucleotide sequence of SVIAK-type clone
66	PPsv-2.pep	translated amino acid sequence of SVIAK CFM
67	PPsv-3	nucleotide sequence of SVIAK-type clone
68	PPsv-3.pep	translated amino acid sequence of SVIAK CFM
69	PPsv-4	nucleotide sequence of SVIAK-type clone
70	PPsv-4.pep	translated amino acid sequence of SVIAK CFM

SEQ	NAME	DESCRIPTION
ID		
NO.	<u> </u>	
71	PPsv-5	nucleotide sequence of SVIAK-type clone
72	PPsv-5.pep	translated amino acid sequence of SVIAK CFM
73	PPsv-6	nucleotide sequence of SVIAK-type clone
74	PPsv-6.pep	translated amino acid sequence of SVIAK CFM
75	Pavsv-A	nucleotide sequence of SVIAK-type clone
76	Pavsv-A.pep	translated amino acid sequence of SVIAK CFM
77	Pavsv-B	nucleotide sequence of SVIAK-type clone
78	Pavsv-B.pep	translated amino acid sequence of SVIAK CFM
79	Pavsv-C	nucleotide sequence of SVIAK-type clone
80	Pavsv-C.pep	translated amino acid sequence of SVIAK CFM
81	RTsv-1	nucleotide sequence of SVIAK-type clone
82	RTsv-1.pep	translated amino acid sequence of SVIAK CFM
83	RTsv-2	nucleotide sequence of SVIAK-type clone
84	RTsv-2.pep	translated amino acid sequence of SVIAK CFM
85	RTsv-3	nucleotide sequence of SVIAK-type clone
86	RTsv-3.pep	translated amino acid sequence of SVIAK CFM
87	Aams-2	nucleotide sequence of (M)SVIAT-type clone
88	Aams-2.pep	translated amino acid sequence of (M)SVIAT CFM
89	Aams-4	nucleotide sequence of (M)SVIAT-type clone
90	Aams-4.pep	translated amino acid sequence of (M)SVIAT CFM
91	Aams-5	nucleotide sequence of SGIAT-type clone
92	Aams-5.pep	translated amino acid sequence of SGIAT CFM
93	Aams-6	nucleotide sequence of (M)SVIAT-type clone
94	Aams-6.pep	translated amino acid sequence of (M)SVIAT CFM
95	Aams-A	nucleotide sequence of (M)SVIAT-type clone
96	Aams-A.pep	translated amino acid sequence of (M)SVIAT CFM
97	Aams-B	nucleotide sequence of (M)SVIAT-type clone
98	Aams-B.pep	translated amino acid sequence of (M)SVIA5 CFM
99	Acams-2	nucleotide sequence of (M)SVIAT-type clone
100	Acams-2.pep	translated amino acid sequence of (M)SVIAT CFM
101	Acams-3	nucleotide sequence of (M)SVIAT-type clone
102	Acams-3.pep	translated amino acid sequence of (M)SVIAT CFM
103	Acams-4	nucleotide sequence of (M)SVIAT-type clone
104	Acams-4.pep	translated amino acid sequence of (M)SVIAT CFM
105	Acams-5	nucleotide sequence of (M)SVIAT-type clone
106 107	Acams-5.pep	translated amino acid sequence of (M)SVIAT CFM
107	Cems-F	nucleotide sequence of (M)SVIAT-type clone
108	Cems-F.pep	translated amino acid sequence of (M)SVIAT CFM
	Cems-G	nucleotide sequence of (M)SVIAT-type clone
110	Cems-G.pep	translated amino acid sequence of (M)SVIAT CFM
111	Cems-H	nucleotide sequence of (M)SVIAT-type clone
112	Cems-H.pep	translated amino acid sequence of (M)SVIAT CFM

	20	
-	'УХ	-

SEQ	NAME	DESCRIPTION
ID	TVA.VIE	PEDCINI ITON
NO.	·	
113	Cems-I	nucleotide sequence of (M)SVIAT-type clone
114	Cems-I.pep	translated amino acid sequence of (M)SVIAT CFM
115	LGAms-5	nucleotide sequence of (M)SVIAT-type clone
116	LGAms-5.pep	translated amino acid sequence of (M)SVIAT CFM
117	LGAms-6	nucleotide sequence of (M)SVIAT-type clone
118	LGAms-6.pep	translated amino acid sequence of (M)SVIAT CFM
119	Mi68Dms	nucleotide sequence of (M)SVIAT-type clone
120	Mi68Dms.pep	translated amino acid sequence of (M)SVIAT CFM
121	Mims-A	nucleotide sequence of (M)SVIAT-type clone
122	Mims-A.pep	translated amino acid sequence of (M)SVIAT CFM
123	Mims-B	nucleotide sequence of (M)SVIAT-type clone
124	Mims-B.pep	translated amino acid sequence of (M)SVIAT CFM
125	Mims-C	nucleotide sequence of (M)SVIAT-type clone
126	Mims-C.pep	translated amino acid sequence of (M)SVIAT CFM
127	PMms-A	nucleotide sequence of (M)SVIAT-type clone
128	PMms-A.pep	translated amino acid sequence of (M)SVIAT CFM
129	PMms-B	nucleotide sequence of (M)SVIAT-type clone
130	PMms-B.pep	translated amino acid sequence of (M)SVIAT CFM
131	PMms-C	nucleotide sequence of (M)SVIAT-type clone
132	PMms-C.pep	translated amino acid sequence of (M)SVIAT CFM
133	PMms-D	nucleotide sequence of (M)SVIAT-type clone
134	PMms-D.pep	translated amino acid sequence of (M)SVIAT CFM
135	PMms-E	nucleotide sequence of (M)SVIAT-type clone
136	PMms-E.pep	translated amino acid sequence of (M)SVIAT CFM
137	PPd57-1ms	nucleotide sequence of (M)SVIAT-type clone
138	PPd57-1ms.pep	translated amino acid sequence of (M)SVIAT CFM
139	PPd57-2ms	nucleotide sequence of (M)SVIAT-type clone
140	PPd57-2ms.pep	translated amino acid sequence of (M)SVIAT CFM
141	PPd57-3	nucleotide sequence of (M)SVIAT-type clone
142	PPd57-3.pep	translated amino acid sequence of (M)SVIAT CFM
143	PPd57-4ms	nucleotide sequence of (M)SVIAT-type clone
144	PPd57-4ms.pep	translated amino acid sequence of (M)SVIAT CFM
145	PPms-1	nucleotide sequence of (M)SVIAT-type clone
146	PPms-1.pep	translated amino acid sequence of (M)SVIAT CFM
147	PPms-2	nucleotide sequence of (M)SVIAT-type clone
148	PPms-2.pep	translated amino acid sequence of (M)SVIAT CFM
149	PPms-E	nucleotide sequence of (M)SVIAT-type clone
150	PPms-E.pep	translated amino acid sequence of (M)SVIAT CFM
151	PPms-G	nucleotide sequence of (M)SVIAT-type clone
152	PPms-G.pep	translated amino acid sequence of (M)SVIAT CFM
153	Pav5ms	nucleotide sequence of (M)SVIAT-type clone
154	Pav5ms.pep	translated amino acid sequence of (M)SVIAT CFM

SEQ	NAME	DESCRIPTION
ID		
NO.		
155	Pavms-2	nucleotide sequence of (M)SVIAT-type clone
156	Pavms-2.pep	translated amino acid sequence of (M)SVIAT CFM
157	Pavms-3	nucleotide sequence of (M)SVIAT-type clone
158	Pavms-3.pep	translated amino acid sequence of (M)SVIAT CFM
159	Pavms-4	nucleotide sequence of (M)SVIAT-type clone
160	Pavms-4.pep	translated amino acid sequence of (M)SVIAT CFM
161	RTms-1	nucleotide sequence of (M)SVIAT-type clone
162	RTms-1.pep	translated amino acid sequence of (M)SVIAT CFM
163	RTms-2	nucleotide sequence of SVSAT-type clone
164	RTms-2.pep	translated amino acid sequence of SVSAT CFM
165	RTms-5	nucleotide sequence of (M)SVIAT-type clone
166	RTms-5.pep	translated amino acid sequence of (M)SVIAT CFM
167	RTms-6	nucleotide sequence of SVIVT-type clone
168	RTms-6.pep	translated amino acid sequence of SVIVT CFM
169	Acasv-B	nucleotide sequence of SVIAK-type clone with a
		stop codon at amino acid position 14
170	Acasv-B.pep	translated amino acid sequence of SVIAK CFM
171	GPd58-1sv	nucleotide sequence of SVIAK-type clone with a
		stop codon at amino acid position 14
172	GPd58-1sv.pep	translated amino acid sequence of SVIAK CFM
173	GPd58-3sv	nucleotide sequence of SVIAK-type clone with a
174	CD160 2	stop codon at amino acid position 14
174 175	GPd58-3sv.pep GPd58-4sv	translated amino acid sequence of SVIAK CFM
1/3	GP038-48V	nucleotide sequence of SVIAK-type clone with a
176	GPd59 deu non	stop codon at amino acid position 14
177	GPd58-4sv.pep Misv-D	translated amino acid sequence of SVIAK CFM nucleotide sequence of SVIAK-type clone with a
1''	מ-۱۷۱۱۶۷	stop codon at amino acid position 14
178	Misv-D.pep	translated amino acid sequence of SVIAK CFM
179	Paysy-D	nucleotide sequence of SVIAK-type clone with a
1//	I dysy-D	stop codon at amino acid position 14
180	Pavsv-D.pep	translated amino acid sequence of SVIAK CFM
181	Aapat-1	amino acid sequence of coral protein disclosed in
	i input i	WO00/46233
182	Aapat-2	amino acid sequence of coral protein disclosed in
- ·		WQ00/46233
183	dT(17)Ad2Ad1	oligonucleotide
184	vispro-F1	oligonucleotide
185	vispro-R1	oligonucleotide
186	pQEprom	oligonucleotide
187	pQErev	oligonucleotide
188	Coral-R1	oligonucleotide

SEQ	NAME	DESCRIPTION
ID		
NO.		
189	A8 (pCGP2918)	nucleotide sequence of full-length cDNA clone
190	A8.aa	translated amino acid sequence of full-length cDNA
		clone
191	D10 (pCGP2920)	nucleotide sequence of full-length cDNA clone
192	D10.aa	translated amino acid sequence of full-length cDNA
100	Fo ( Garage )	clone
193	S3 (pCGP2924)	nucleotide sequence of full-length cDNA clone
194	S3.aa	translated amino acid sequence of full-length cDNA clone
195	T3 (pCGP2922)	nucleotide sequence of full-length cDNA clone
196	T3.aa	translated amino acid sequence of full-length cDNA clone
197	D1 (pCGP2919)	nucleotide sequence of full-length cDNA clone
198	D1.aa	translated amino acid sequence of full-length cDNA
		clone
199	S1 (pCGP2923)	nucleotide sequence of full-length cDNA clone
200	S1.aa	translated amino acid sequence of full-length cDNA clone
201	T1 (pCGP2921)	nucleotide sequence of full-length cDNA clone
202	T1.aa	translated amino acid sequence of full-length cDNA
		clone
203	Kpn.6His.F	oligonucleotide
204	T1/A8.Sal.R	oligonucleotide
205	TSSU-Fnew	oligonucleotide
206	TSSU-R	oligonucleotide
207	AscI-ER.F	oligonucleotide
208	ER-BamHI.R	oligonucleotide
209	CP-HDEL-PacLR	oligonucleotide
210	Pst-mGFP4F	oligonucleotide
211	mGFP4-PacIR	oligonucleotide
212	visproF1-new	oligonucleotide
213	MSV-RBS	oligonucleotide
214	SVIAK-RBS	oligonucleotide
215	POC 220 H6	oligonucleotide
216	Rtms-5v	mutated variant amino acid sequence from Rtms-5
217	gtCP	(SEQ ID NO:166)
218	poc4	translated amino acid sequence of SVIAK CFM
219	baspoc3	translated amino acid sequence of SVIAK CFM
220	dsFP593	translated amino acid sequence of SVIAK CFM
221	drFP583, also known as	translated amino acid sequence of a CFM
221	dsRed583	translated amino acid sequence of a CFM

- 31 -

SEQ ID NO.	NAME	DESCRIPTION
222	gfp	translated amino acid sequence of a CFM
223	MGFP-4	nucleotide sequence from GFP-4 from Aequorea
		victoria (jelly fish), mutated for plants
224	MGFP-4.pep	translated amino acid sequence of GFP-4 CFM
225	BASPOC4	translated amino acid sequence of a CFM
226	AsFP595	translated amino acid sequence of a CFM
227	TICS-His-FWD	oligonucleotide
228	TICS-His-REV	oligonucleotide
229	mGFP4-HDEL-PacR	oligonucleotide
230	T1.N-QN(AAT)SQ(CAG)	oligonucleotide
231	T1.S-IS(TCC)>I(ATC)	oligonucleotide
232	YGFP3UP	oligonucleotide
233	YGFP3DO	oligonucleotide
234	RFPUP1	oligonucleotide
235	RFPDO1	oligonucleotide
236	MSVIATUP	oligonucleotide
237	COFPDO	oligonucleotide
238	ATP4PROMUP2	oligonucleotide
239	ATP4DO2	oligonucleotide
240	ATP7TUP	oligonucleotide
241	ATP7TDO	oligonucleotide
242	SPPDYTLEFP	N-terminal amino acid sequence of a CFM
243	SPPDYTLERP	N-terminal amino acid sequence of a CFM
244	(D)SS(P)E	N-terminal amino acid sequence of a CFM
245	SYLPN	N-terminal amino acid sequence of a CFM
246	SYLQN	N-terminal amino acid sequence of a CFM
247	MEGIVNG-A	oligonucleotide
248	MEGIVNG-T	oligonucleotide
249	MEGIVNG-C	oligonucleotide
250	REV-MEG-T	oligonucleotide
251	REV-MEG-C	oligonucleotide

- 32 -

#### DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

The present invention is predicated on the identification of peptides, polypeptides and proteins having one or more amino acid sequences or one or more amino acid sequences which exhibit color-facilitating properties, either on their own or following interaction with one or more other amino acids and nucleic acid moleclues encoding same. Such peptides, polypeptides and proteins are referred to herein as "color-facilitating molecules" or "CFMs". The present invention contemplates a range of uses of CFMs, including their use as color expression markers and as color intensifiers, as well as in gel-like formulations for use as photon traps and in light-filtering formulations such as topically-applied sun creens.

The present invention further contemplates the use of genetic material encoding CFMs to generate eukaryotic or prokaryotic cells or eukaryotic or prokaryotic cell tissue which, in the presence of the CFMs, exhibit altered visual characteristics to the human eye in the absence of excitation of the CFMs by extraneous non-white light or particle emission.

Such altered visual characteristics are also referred to as being altered to the naked, unaided eye. Reference to "naked" or "unaided" is not to imply that the eye may not require magnification aids such as in the form of spectacles or glasses or a magnifying glass. Reference to extraneous light or particle emission includes ultraviolet (UV) light, blue laser light, plasma irradiation,  $\gamma$ -irradiation, particle irradiation, single wavelength light such as 340 nm, 382 nm, 396 nm, 405 nm, 475 nm, 490 nm, 575 nm or other forms of emission or particle bombardment. It does not include white light.

Accordingly, one aspect of the present invention provides an isolated nucleic acid molecule comprising a nucleotide sequence encoding a color-facilitating molecule (CFM) which, in a cell, alone or together with one or more other molecules imparts an altered visual characteristic to said cell when visualized by a human eye in the absence of excitation by extraneous non-white light or particle emission.

10

15

20

- 33 -

Preferably, the nucleic acid molecule is derived from Anemonia majano, Anemonia sulcata, Clavularia sp, Zoanthus sp, Discosoma sp (e.g. Discosoma striata), Aequorea sp (e.g. Aequorea victoria), Anthozoa sp, Cassiopea sp, (e.g. Cassiopea xamachana), Millepora sp, Acropora sp (e.g. Acropora aspera and Acropora nobilis), Montipora sp, Porites murrayensis, Pocillopora damicormis, Pavona descussaca, Acanthastrea sp, Platygyra sp or Caulastrea sp.

In a preferred embodiment, the nucleic acid molecule encodes a CFM with an amino acid at its N-terminal region selected from SVIAK (SEQ ID NO:5), (M)SVIAT (SEQ ID NO:6), SGIAT (SEQ ID NO:7), SVIVT (SEQ ID NO:8) or SVSAT (SEQ ID NO:9). Even more particularly, the CFM comprises an amino acid sequence selected from SVIAT QMTY KVYM SGT (SEQ ID NO:10), SVIAT QMTY KVYM PGT (SEQ ID NO:11), SVIAT QVTY KVYM SGT (SEQ ID NO:12), SGIAT QMTY KVYM SGT (SEQ ID NO:13), SVIVT QMTY KVYM SGT (SEQ ID NO:14), SVSAT QMTY KVYM SGT (SEQ ID NO:15), SVIAK QMTY KVYM SGT (SEQ ID NO:16), SVIAK QMTY KVYM SGT (SEQ ID NO:16), SVIAK QMTY KVYM SDT (SEQ ID NO:17) and SVIAK QMTY X<sub>1</sub>X<sub>2</sub>YX<sub>3</sub> SGT (SEQ ID NO:18) wherein X<sub>1</sub>, X<sub>2</sub> and X<sub>3</sub> may be any amino acid provided that X<sub>1</sub> is not K; X<sub>2</sub> is not V; X<sub>3</sub> is not M.

In a particular embodiment, the present invention provides an isolated nucleic acid molecule comprising a nucleotide sequence encoding a CFM or a fragment, variant or derivative thereof, wherein said isolated nucleic acid molecule comprises a nucleotide sequence selected from the group consisting of SEQ ID NOs:19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, 131, 133, 135, 137, 139, 141, 143, 145, 147, 149, 151, 153, 155, 157, 159, 161, 163, 165, 167, 169, 171, 173, 175, 177, 179, 189, 191, 193, 195, 197, 199 and 201, or a biologically active fragment or derivative of these.

Particular preferred nucleic acid molecules comprise the nucleotide sequences set forth in SEQ ID NOs:189, 191, 193, 195, 197, 199 and 201.

- 34 -

The nucleic acid molecule is regarded as genetic material and generally comprises a coding region encoding a CFM optionally operably linked to a single or multiple promoters. In one embodiment, the nucleic acid molecule is a genetic construct under the control of (i.e. operably linked to) a single promoter. In another embodiment, the genetic construct is a bicistronic, tricistronic or multicistronic construct carrying the gene encoding the CFM and optionally other genes such as encoding a reporter molecule.

As used herein, the terms "nucleic acid molecule" including "genetic material" refers to any single-stranded or double-stranded nucleic acid molecule which at least comprises deoxyribonucleotides and/or ribonucleotides, including DNA (cDNA or genomic DNA), RNA, mRNA, or tRNA, amongst others. The combination of such molecules with non-nucleotide substituents derived from synthetic means or naturally-occurring sources is also contemplated by the present invention. Genetic material may also include sequences optimized for expression of codons in a particular host cell.

15

20

25

The present invention extends to derivatives of the nucleic acid molecules and such derivatives includes any isolated nucleic acid molecule which comprises at least 10 and preferably at least 20 contiguous nucleotides derived from the genetic sequence as described herein according to any embodiment. A derivative includes a part, fragment, portion or analog. A derivative also includes a fusion molecule between two or more genetic sequences encoding CFMs.

The present invention also comprises analogs of the nucleic acid molecules. An "analog" means any isolated nucleic acid molecule which is substantially the same as a nucleic acid molecule of the present invention or its complementary nucleotide sequence as described herein according to any embodiment, notwithstanding the occurrence of any non-nucleotide constituents not normally present in said isolated nucleic acid molecule, for example carbohydrates, radiochemicals including radionucleotides, reporter molecules such as, but not limited to, alkaline phosphatase or horseradish peroxidase, amongst others.

A "homolog" is a functionally similar molecule from a different species or strain.

- 35 -

Generally, analogs or derivatives of the nucleic acid molecule of the invention are produced by synthetic means or alternatively, derived from naturally-occurring sources. For example, the nucleotide sequence of the present invention may be subjected to mutagenesis to produce single or multiple nucleotide substitutions, deletions and/or insertions. A derivative encompasses a nucleotide sequence modified for optimized or enhanced codon usage in a particular cell.

5

10

15

20

25

The genetic sequence of the present invention may comprise a sequence of nucleotides or be complementary to a sequence of nucleotides which comprise one or more of the following: a promoter sequence, a 5' non-coding region, a cis-regulatory region such as a functional binding site for transcriptional regulatory protein or translational regulatory protein, an upstream activator sequence, an enhancer element, a silencer element, a TATA box motif, a CCAAT box motif, or an upstream open reading frame, transcriptional start site, translational start site, and/or nucleotide sequence which encodes a leader sequence. The genetic sequence also encodes the CFM.

The term "5" non-coding region" is used herein in its broadest context to include all nucleotide sequences which are derived from the upstream region of an expressible gene, other than those sequences which encode amino acid residues which comprise the polypeptide product of said gene, wherein 5' non-coding region confers or activates or otherwise facilitates, at least in part, expression of the gene.

The nucleic acid molecule may also be regarded as a gene. The term "gene" is used in its broadest context to include both a genomic DNA region corresponding to the gene as well as a cDNA sequence corresponding to exons or a recombinant molecule engineered to encode a functional form of a product. The term "gene" is used in its broadest sense and includes cDNA corresponding to the exons of a gene. Accordingly, reference herein to a "gene" is to be taken to include:-

- 36 -

- (i) a classical genomic gene consisting of transcriptional and/or translational regulatory sequences and/or a coding region and/or non-translated sequences (i.e. introns, 5'- and 3'- untranslated sequences); or
- 5 (ii) mRNA or cDNA corresponding to the coding regions (i.e. exons) and 5'- and 3'untranslated sequences of the gene.

The term "gene" is also used to describe synthetic or fusion molecules encoding all or part of a functional product.

10

As used herein, the term "cis-acting sequence" or "cis-regulatory region" or similar term shall be taken to mean any sequence of nucleotides which is derived from an expressible genetic sequence wherein the expression of the first genetic sequence is regulated, at least in part, by said sequence of nucleotides. Those skilled in the art will be aware that a cis-regulatory region may be capable of activating, silencing, enhancing, repressing or otherwise altering the level of expression and/or cell-type-specificity and/or developmental specificity of any structural gene sequence.

Reference herein to a "promoter" is to be taken in its broadest context and includes the transcriptional regulatory sequences of a classical genomic gene, including the TATA box which is required for accurate transcription initiation, with or without a CCAAT box sequence and additional regulatory elements (i.e. upstream activating sequences, enhancers and silencers) which alter gene expression in response to developmental and/or environmental stimuli, or in a tissue-specific or cell-type-specific manner. A promoter is usually, but not necessarily, positioned upstream or 5', of a structural gene, the expression of which it regulates. Furthermore, the regulatory elements comprising a promoter are usually positioned within 2 kb of the start site of transcription of the gene.

In the present context, the term "promoter" is also used to describe a synthetic or fusion molecule, or derivative which confers, activates or enhances expression of a structural gene or other nucleic acid molecule, in a plant cell. Preferred promoters according to the

- 37 -

subject invention may contain additional copies of one or more specific regulatory elements to further enhance expression in a cell, and/or to alter the timing of expression of a structural gene to which it is operably connected.

In a preferred embodiment, the nucleic acid molecules are expressed in a cell. The cell may be a eukaryotic or prokaryotic cell. Reference to a eukaryotic cell includes a mammalian animal cell, a non-mammalian animal cell or a plant cell. Insofar as the eukaryotic cell is a plant cell, the plant cell may be part of a plant callus or a whole plant. Reference to a "plant" includes ornamental or flowering plants or parts thereof such as flowers, roots, leaves, stems, seeds, fruit or fibers. Particularly preferred plant cells are those selected from rose, carnation, lisianthus, petunia, lily, tulip, pansy, gerbera or chrysanthemum.

The CFM is preferably a GFP or a derivative or homolog thereof such as a non-fluorescent GFP homolog.

15

Another aspect of the present invention provides an isolated color-facilitating molecule (CFM) comprising a polypeptide which, in a cell, alone or together with one or more other molecules imparts an altered visual characteristic to said cell when visualized by a human eye in the absence of excitation by extraneous non-white light or particle emission.

20

25

30

The CFM of the present invention is preferably a protein comprising a sequence of amino acids such that upon folding, the sequence alone or following interaction with one or more other amino acids which may be within the same molecule or in another molecule such as in a dimer, trimer or oligomer exhibits chromophore or fluorophore properties. Particularly useful proteins comprise the contiguous amino acid sequence Gln-Tyr-Gly (QYG). Even more preferably, the protein is a GFP or a homolog or derivative thereof. An example of a homolog of a GFP is a non-fluorescent GFP homolog. An example of a derivative of a GFP or non-fluorescent GFP homolog is a GFP modified to cause a shift in the ratio of excitation or emission peaks. Such modifications may result in a more intense fluorescence or may exhibit altered or weaker fluorescence.

- 38 -

Any number of GFP or non-fluorescent GFP homologs or other derivatives may be employed as CFMs in accordance with the present invention. Examples of such molecules are from Anemonia majano, Anemonia sulcata, Clavularia sp, Zoanthus sp, Discosoma sp (e.g. Discosoma striata), Aequorea sp (e.g. Aequorea victoria), Anthozoa sp, Cassiopea sp, (e.g. Cassiopea xamachana), Millepora sp, Acropora sp (e.g. Acropora aspera and Acropora nobilis), Montipora sp, Porites murrayensis, Pocillopora damicormis, Pavona descussaca, Acanthastrea sp, Platygyra sp and Caulastrea sp.

Particularly preferred protein sequences which constitute CFMs of the present invention comprise one of the following sequences of amino acids towards the amino-terminal end of the polypeptide: "SVIAK" (SEQ ID NO:5), "(M)SVIAT" (SEQ ID NO:6), "SGIAT" (SEQ ID NO:7), "SVIVT" (SEQ ID NO:8), or "SVSAT" (SEQ ID NO:9).

Examples of such preferred protein sequences may be selected from the group consisting of:

SVIAT QMTY KVYM SGT (SEQ ID NO:10);

SVIAT QMTY KVYM PGT (SEQ ID NO:11);

SVIAT QVTY KVYM SGT (SEQ ID NO:12);

20 SGIAT QMTY KVYM SGT (SEQ ID NO:13);

SVIVT QMTY KVYM SGT (SEQ ID NO:14);

SVSAT QMTY KVYM SGT (SEQ ID NO:15);

SVIAK QMTY KVNM SGT (SEQ ID NO:16);

SVIAK QMTY KVYM SDT (SEQ ID NO:17); and

25 SVIAK QMTY X<sub>1</sub>X<sub>2</sub>YX<sub>3</sub> SGT (SEQ ID NO:18),

wherein  $X_1$ ,  $X_2$  and  $X_3$  may be any amino acid provided that  $X_1$  is not  $X_2$  is not  $X_3$  is not  $X_4$ .

30 Accordingly, in another aspect of the present invention there is provided an isolated polypeptide, or a biologically active fragment thereof, or a variant or derivative of these,

- 39 -

said polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NOs:10, 11, 12, 13, 14, 15, 16, 17 and 18, with the proviso that, in said isolated polypeptide or biologically active fragment or variant or derivative of SEQ ID NO:18,  $X_1$  is not lysine,  $X_2$  is not valine, and  $X_3$  is not methionine.

5

10

15

20

25

Particularly suitable molecules comprise an amino acid sequence selected from the group consisting of SEQ ID NOs:20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178, 180, 190, 192, 194, 196, 198, 200 and 202.

Accordingly, a preferred embodiment of the present invention provides an isolated polypeptide, or a biologically active fragment thereof, or a variant or derivative of these, said polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NOs:20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178, 180, 190, 192, 194, 196, 198, 200 and 202 provided that, where the said biologically active fragment or variant or derivative comprises the sequence SVIAK QMTY X<sub>1</sub>X<sub>2</sub>YX<sub>3</sub> SGT, X<sub>1</sub> is not lysine, X<sub>2</sub> is not valine, and X<sub>3</sub> is not methionine.

Such isolated polypeptides, when present in a prokaryotic or eukaryotic cell or group of prokaryotic or eukaryotic cells such as in plant cells in the form of plant tissue or plant callus, may alone or in combination with one or more other molecules impart an altered visual characteristic to said cell or group of cells when visualized by a human eye in the absence of excitation by extraneous non-white light or particle emission.

Accordingly, another aspect of the present invention provides a prokaryotic or eukaryotic cell or group of prokaryotic or eukaryotic cells in the form of tissue wherein said cell or

- 40 -

group of cells or their parent cells are genetically modified to enable the production of a color-facilitating molecule (CFM) which alone or together with one or more other molecules imparts an altered visual characteristic to said cell or group of cells when visualized by a human eye in the absence of excitation by extraneous non-white light or particle emission.

The CFM is as herein defined and in a preferred embodiment includes polypeptides having amino acid sequence selected from the list comprising SEQ ID NOs:20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178, 180, 190, 192, 194, 196, 198, 200 and 202 provided that, where the said amino acid sequence comprises the sequence SVIAK QMTY X<sub>1</sub>X<sub>2</sub>YX<sub>3</sub> SGT, X<sub>1</sub> is not lysine, X<sub>2</sub> is not valine, and X<sub>3</sub> is not methionine.

15

20

25

30

A "eukaryotic" cell is regarded as any cell which is not characterized as being a "prokaryotic" cell. Particularly useful eukaryotic cells are plant cells as well as fungi and yeast. Other eukaryotic cells, however, are also contemplated such as mammalian cells, non-mammalian animal cells including insect cells as well as plant cells. A "plant" may be regarded as a monocotyledonous or dicotyledonous plant and includes ornamental and crop plants. Reference to "tissue" includes plant callus. A "prokaryotic cell" is generally a cell comprising a nucleus not surrounded by a nuclear membrane and includes bacteria and microbial cells. Such prokaryotic cells include *Pseudomonas* sp., *E. coli, Enterobacter* sp., *Salmonella* sp., *Klebsiella* sp., *Acetobacter* sp., *Staphylocous* sp., *Streptococcus* sp. or *Bacillus* sp., amongst many others.

In a preferred embodiment, the present invention provides a plant cell or group of plant cells such as in the form of plant tissue or plant callus wherein said plant cells or group of plant cells or their parent cells are genetically modified to enable production of a CFM which alone or in combination with one or more other molecules imparts an altered visual

- 41 -

characteristic to said cell or group of cells when visualized by a human eye in the absence of excitation by extraneous non-white light or particle emission.

Particularly preferred plants are ornamental and flowering plants. Particularly useful plants contemplated by the present invention include but are not limited to rose, carnation, lisianthus, petunia, lily, tulip, pansy, gerbera and chrysanthemum.

Reference herein to a "plant" includes parts of plants. Similarly, reference herein to "plant tissue" includes parts of plants. Examples of such plant parts, include but are not limited to, flowers, roots, leaves, stems, seeds, fruit and fibres. The term "flowers" includes parts of flowers such as petals, petioles, flower heads and flower buds. Plant tissue may also include callus material as well as embryogenic or non-embryogenic material. The term "fibre" includes cotton and hemp fibres.

Accordingly, another aspect of the present invention is directed to a plant or part of a plant including a flower, root, leaf, stem, seed, fruit or fibre or reproductive portion of said plant or cells of said plant wherein said plant or plant part comprises cells genetically modified to enable production of a CFM which alone or in combination with one or other molecules imparts an altered visual characteristic to said cells when visualized by a human eye in the absence of excitation by extraneous non-white light or particle emission.

The term "genetically modified" is used in its broadest sense and includes introducing gene(s) into cells, mutating gene(s) in cells and altering or modulating the regulation of gene(s) in cells.

25

A "part" of a plant includes flowers (e.g. cut or severed flowers), petals, stems, leaves and fibrous material such as cotton and vegetative, propagative and reproductive material such as cuttings, pollen, seeds and callus.

30 The altered visual characteristic may be exhibited by all cells in the plant or in selected tissue or plant parts such as flowers, roots, leaves, stems, seeds, fruit or fibres. The

- 42 -

production of the CFM may, therefore, be tissue specific or tissue preferential. Furthermore, CFM production may be developmentally dependent, determined, influenced or otherwise regulated.

The CFM may be produced in the whole plant with the use of a constitutive promoter such as cauliflower mosaic virus (CaMV) 35S promoter, operably connected or operably linked to a gene or other nucleic acid molecule encoding the CFM. Alternatively, the molecule may be confined to, for example, petal tissue, epidermal cell layers of petals or to different organelles within cells. For example, a floral specific promoter such as a chalcone synthase promoter substantially limits a colored protein expression to flower petals.

The use of some gene promoters (e.g. 35S) may produce CFM accumulation in the cytoplasm of transformed cells and confer a visible color to the plant tissue. The CFM may be targeted to different organelles within the plant cell to confer a color change in the tissue visible to the naked unaided eye under white light illumination. The CFM can be targeted to plastids using a chloroplast transit peptide fused in-frame with the colored protein cDNA sequence. An example of a plastid transit peptide that can be used is the transit peptide of the small subunit of ribulose-1, 5-bisphosphate-carboxylase (e.g. InCheol et al., Molecular Breeding 5: 453-461, 1999). The targeting of a CFM to plastids can dramatically increase the total amount of protein accumulated (InCheol et al., 1999, supra) and thereby increase color intensity.

15

20

25

Chromoplasts are numerous in the petals of some flowers, leaves and fruit. A chromoplast specific transit peptide fused in-frame with the protein cDNA sequence may be used to modify flower or other tissue color with a much reduced potential for interfering with total plant photosynthetic activity, as may occur if an constitutive promoter and a chloroplast transit peptide were used to target the CFM. The use of a chromoplast transit peptide and a floral specific promoter may be optimal for the modification of flower color.

30 It may be beneficial to target all CFMs to the vacuole or endoplasmic reticulum to avoid any detrimental effects to the transformed cells or plants. An example of an endoplasmic

- 43 -

reticulum targeting peptide sequence that can be used is the amino acid sequence HDEL (Haseloff et al., 1997, supra). The CFM may also be targeted to the cell wall.

5

10

15

20

The term "operably connected" or "operably linked" in the present context means placing a structural gene (e.g. a nucleic acid molecule encoding a CFM) under the regulatory control of a promoter which then controls expression of the gene. Promoters and the like are generally positioned 5' (upstream) to the genes which they control. In the construction of heterologous promoter/structural gene combinations, it is generally preferred to position the genetic sequence or promoter at a distance from the gene transcription start site that is approximately the same as the distance between that genetic sequence or promoter and the gene it controls in its natural setting, i.e., the gene from which the genetic sequence or promoter is derived. As is known in the art, some variation in this distance can be accommodated without loss of function. Similarly, the preferred positioning of a regulatory sequence element with respect to a heterologous gene to be placed under its control is defined by the positioning of the element in its natural setting, i.e., the genes from which it is derived.

The cells genetically modified to enable production of a CFM may be the cells into which genetic material has been introduced or they may represent progeny of genetically modified parent cells.

Accordingly, the present invention contemplates a method for generating a transgenic plant or part of a plant, wherein said plant or plant part comprises cells genetically modified to enable production of a CFM which alone or in combination with one or other molecules imparts an altered visual characteristic to said cells when visualized by a human eye in the absence of excitation by extraneous non-white light or particle emission, said method comprising introducing into said cells an isolated nucleic acid molecule encoding said CFM.

30 Preferably, the CFM is derived from Anemonia majano, Anemonia sulcata, Clavularia sp, Zoanthus sp, Discosoma sp (e.g. Discosoma striata), Aequorea sp (e.g. Aequorea victoria),

- 44 -

Anthozoa sp, Cassiopea sp, (e.g. Cassiopea xamachana), Millepora sp, Acropora sp (e.g. Acropora aspera and Acropora nobilis), Montipora sp, Porites murrayensis, Pocillopora damicormis, Pavona descussaca, Acanthastrea sp, Platygyra sp or Caulastrea sp.

More preferably, the CFM comprises an amino acid sequence in its N-terminal end selected from SVIAK (SEQ ID NO:5), (M)SVIAT (SEQ ID NO:6), SGIAT (SEQ ID NO:7), SVIVT (SEQ ID NO:8) or SVSAT (SEQ ID NO:9).

Even more preferably, the CFM comprises an amino acid sequence selected from the list comprising SVIAT QMTY KVYM SGT (SEQ ID NO:10), SVIAT QMTY KVYM PGT (SEQ ID NO:11), SVIAT QVTY KVYM SGT (SEQ ID NO:12), SGIAT QMTY KVYM SGT (SEQ ID NO:13), SVIVT QMTY KVYM SGT (SEQ ID NO:14), SVSAT QMTY KVYM SGT (SEQ ID NO:15), SVIAK QMTY KVNM SGT (SEQ ID NO:16), SVIAK QMTY KVYM SDT (SEQ ID NO:17) and SVIAK QMTY X<sub>1</sub>X<sub>2</sub>YX<sub>3</sub> SGT (SEQ ID NO:18) wherein X<sub>1</sub>, X<sub>2</sub> and X<sub>3</sub> may be any amino acid provided that X<sub>1</sub> is not K; X<sub>2</sub> is not V; X<sub>3</sub> is not M.

Most preferably, the CFM is encoded by a nucleotide sequence selected from the list comprising SEQ ID NOs:19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, 131, 133, 135, 137, 139, 141, 143, 145, 147, 149, 151, 153, 155, 157, 159, 161, 163, 165, 167, 169, 171, 173, 175, 177, 179, 189, 191, 193, 195, 197, 199 and 201.

Another aspect of the present invention provides a transgenic plant wherein said plant or a part thereof such as a flower, leaf, root, stem, seed, fruit or fibre exhibits an altered visual characteristic to a human eye in the absence of extraneous non-white light or particle emission wherein cells of said transgenic plant or of a parent plant have been genetically modified to enable production of a CFM.

10

20

- 45 -

As stated above, the present invention extends to genetically modified mammalian cells, non-mammalian animal cells as well as plant cells.

Farmers use conventional breeding techniques to develop new colors in animals and animal products for the market, for example, colored wools and leathers or hides. Presently the main way of coloring these products toobtain novel colors is by using dyes or tints or paints or pigments on natural colored products. However, the use of the CFMs of the present invention can be employed to produce a transgenic animal which exhibits a novel color: for example, sheep with blue or red colored fleece, cows with red colored hide.

10

15

Specifically the CFM can be used in a range of agriculturally important animals such as but not limited to sheep, pigs, cattle, horses, goats, llamas, fish, ostriches, emus, ducks and chickens. Accordingly, another aspect of the present invention provides a transgenic mammalian or non-mammalian animal cell or transgenic non-human mammal or non-mammalian animal comprising said cells, said cells exhibiting an altered visual characteristic to a human eye in the absence of extraneous non-white light or particle emission wherein cells of said transgenic plant, mammal or animal or plant cells thereof have been genetically modified to enable production of a CFM.

20 The CFM is as herein defined. Production of the CFM may be constitutive or developmental or may be inducible in response to internal or external stimulus including stress.

Reference herein to a "color-facilitating molecule", "CFM", "protein", "GFP" or "non-fluorescent GFP-homolog" includes fragments, derivatives, variants and homologs thereto. Examples of derivatives include mutants, parts, fragments and portions of these molecules including single or multiple amino acid substitutions, deletions and/or additions to the molecules. Derivatives also include fusion molecules between two or more CFMs or between a CFM and another molecule such as a leader sequence, targeting sequence, expression-facilitating sequence and/or a reporter molecule capable of providing an identifiable signal.

- 46 -

As stated above a derivative also includes a modified form providing altered ratios of excitation or emission spectra. In addition, or as a consequence of the altered ratios of excitation or emission, the modified GFP or their homologs may have a more intense color-producing capacity relative to an unmodified molecule.

5

10

15

20

Furthermore, other proteins may be used in conjunction with the CFMs to alter the visual characteristics of the cells. Examples of other proteins include copper containing proteins containing one or more type 1 (CuII) motifs as found in the Fet3 protein from Saccharomyces cerevisiae (Hassett et al., Journal of Biological Chemistry 273: 23274 -23282, 1998) and other multinuclear copper ferroxidase enzymes such as laccase. ceruloplasmin and ascorbate oxidase (Messerschmidt and Huber, Eur. J. Biochem. 187: 341 - 352, 1990). Similarly, the mononuclear blue or type 1 copper proteins (cupredoxins), such as plastocyanin, azurin, pseudoazurin, plantacyanin, rusticyanin, amicyanin, auracyanin and halocyanin (Nersissian et al., Protein Science 5: 2184 - 2192, 1996). These proteins have not been associated with pigmentation in nature. However, when these proteins are concentrated an intense blue color is evident (Hassett et al., 1998, supra; Messerschmidt and Huber, 1990, supra). The over-expression of a type 1 (CuII) containing protein in flowers and other plant tissues under conditions that allow correct folding and acquisition of Cu ions can modify or impart a color visible to the naked unaided eye under white light. Reference to "in conjunction" includes reference to a fusion protein between a CFM and another protein such as a cuproprotein and well as the expression of nucleotide sequences in multicistronic form encoding a CFM and at least one other protein.

Another aspect of the present invention provides a eukaryotic or prokaryotic cell or a group of eukaryotic or prokaryotic cells in the form of a tissue wherein said cell or group of cells or their parent cells are genetically modified to produce a GFP or derivative or homolog thereof such as a non-fluorescent GFP homolog which imparts an altered visual characteristic on said cell or group of cells when visualized by a human eye in the absence of excitation by extraneous non-white light or particle emission.

- 47 -

Preferably, the eukaryotic cells are plant cells or plant tissue. The eukaryotic cells may, however, be mammalian cells or non-mammalian animal cells. Reference to "plant tissue" includes "callus".

Accordingly, another aspect of the present invention is directed to a plant or part of a plant including a flower, root, leaf, stem, seed, fruit or fibre or reproductive portion of said plant or cells of said plant wherein said plant or plant part comprises cells genetically modified to enable production of a GFP or a derivative or homolog thereof such as a non-fluorescent GFP homolog which imparts an altered visual characteristic to said cells when visualized by a human eye in the absence of excitation by extraneous non-white light or particle emission.

A particularly preferred embodiment the present invention is directed to a plant or part of a plant including a flower, root, leaf, stem, seed, fruit or fibre or reproductive portion of said plant or cells of said plant wherein said plant or plant part comprises cells genetically modified to comprise a polynucleotide comprising the nucleotide sequence set forth in any one of SEQ ID NOs:19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, 131, 133, 135, 137, 139, 141, 143, 145, 147, 149, 151, 153, 155, 157, 159, 161, 163, 165, 167, 169, 171, 173, 175, 177, 179, 189, 191, 193, 195, 197, 199 or 201, or a derivative or homolog of any of these, thereby enabling production of a CFM which alone or in combination with one or more other molecules imparts an altered visual characteristic to said cell or group of cells when visualized by a human eye in the absence of excitation by extraneous non-white light or particle emission.

20

25

30

The present invention particularly provides, in a preferred embodiment, a genetically modified plant carrying flowers having an altered flower color relative to a non-genetically modified plant as well as cut flowers from such a plant. Reference herein to a "genetically modified plant" includes progeny of a genetically modified plant as well as hybrids and derivatives of a genetically modified plant.

- 48 -

The altered coloration of eukaryotic cells such as plant cells is useful not only for the ornamental plant market but also as propriety tags, for example, of seeds, root stock, flowers, crops and whole plants and plant parts. This may be particularly important for distinguishing between transgenic and non-transgenic crops, plants and other horticultural products. Furthermore, the modification of visible color in cotton fibre or hemp is a useful means of reducing the toxicity of dye processes in color fabric manufacture. The modification of visible color in edible and/or ornamental fungal species may also be used to differentiate and enhance marketability.

10

30

The modification of visible color in fruit and vegetables may be used to differentiate and enhance their marketability. A suitable gene promoter may be used to control the expression of the CFM to signal optimal time to, for example, harvest crop plants including harvesting plant parts such as flowers or seeds. In addition, a stress-inducible promoter may be utilized to promote an early warning of water or pathogen stress, allowing for early intervention by the grower and subsequent reduction in economic loss.

Other uses for the CFM of the present invention include, for example, the production of novel colored plant extracts wherein the extract includes, for example, a flavouring or food additive or health product or beverage or juice or coloring. Beverages may include but are not limited to wines, spirits, beers, teas, coffee, milk and dairy products.

The CFM may be used to alter the color of many products such as but not limited to foods (e.g. breads and yeast products, confectionery), beverages (see above) or novelty items 25 (e.g. toys).

A further aspect of the present invention provides a transfected or transformed cell, tissue, organ or non-cellular material which contains or is capable of producing a CFM or a functional derivative or homolog thereof. Preferably, the CFM is a protein such as GFP or a non-fluorescent GFP-homolog.

- 49 -

The genetic construct(s) of the present invention may be introduced into a cell by various techniques known to those skilled in the art. The technique used may vary depending on the known successful techniques for that particular organism.

Techniques for introducing recombinant DNA into cells include, but are not limited to, transformation using CaCl<sub>2</sub> and variations thereof, direct DNA uptake into protoplasts, PEG-mediated uptake to protoplasts, microparticle bombardment, electroporation, microinjection of DNA, microparticle bombardment of tissue explants or cells, vacuum-infiltration of tissue with nucleic acid, and T-DNA-mediated transfer from Agrobacterium to the plant tissue.

For microparticle bombardment of cells, a microparticle is propelled into a cell to produce a transformed cell. Any suitable ballistic cell transformation methodology and apparatus can be used in performing the present invention. Exemplary apparatus and procedures are disclosed by Stomp *et al.* (U.S. Patent No. 5,122,466) and Sanford and Wolf (U.S. Patent No. 4,945,050). When using ballistic transformation procedures, the genetic construct may incorporate a plasmid capable of replicating in the cell to be transformed.

Examples of microparticles suitable for use in such systems include 0.1 to 10 μm and more particularly 10.5 to 5 μm tungsten or gold spheres. The DNA construct may be deposited on the microparticle by any suitable technique, such as by precipitation.

Once introduced into cells such as plant tissue, the expression of a CFM may be assayed in a transient expression system or it may be determined after selection for stable integration within for example, the plant genome. Hence, a CFM of the present invention may be useful as an expression marker. For example, genetic material encoding a CFM of the present invention, optionally operably linked to a single or multiple promoters, may be introduced into cells as a fluorescent "tag", optionally fused with one or more other nucleic acid sequences that may encode a polypeptide or a regulatory nucleotide sequence. In this manner, a CFM fused with another polypeptide may be useful in assessing subcellular

30

- 50 -

localisation of the fusion or, alternatively, as an expression marker for assessing possible activity of the regulatory nucleotide sequence in a given host cell.

Host cells may be prokaryotic cells, for example bacterial, or eukaryotic cells, for example yeast, plant, and animal cells, including human. Preferred host cells are bacterial or plant.

Plant tissue capable of subsequent clonal propagation, whether by organogenesis or embryogenesis, may be transformed with a genetic construct of the present invention and a whole plant generated therefrom. The particular tissue chosen will vary depending on the clonal propagation systems available for, and best suited to, the particular species being transformed. Exemplary tissue targets include leaf disks, pollen, embryos, cotyledons, hypocotyls, megagametophytes, callus tissue, existing meristematic tissue (e.g. apical meristem, axillary buds, and root meristems), and induced meristem tissue (e.g. cotyledon meristem and hypocotyl meristem).

15

10

١

The regenerated transformed plants may be propagated by a variety of means, such as by clonal propagation or classical breeding techniques. For example, a first generation (or T1) transformed plant may be selfed to give homozygous second generation (or T2) transformant, and the T2 plants further propagated through classical breeding techniques.

20

25

Any number of GFP or non-fluorescent GFP-homologs may be employed provided that the GFP or its homolog or other CFM imparts on a cell or group of cells an altered visual characteristic to the human eye in the absence of extraneous non-white light or particle emission. Examples of CFMs contemplated herein include but are not limited to non-fluorescent GFP-homologs such as that encoded by asFP595 (Lukyanov et al., 2000, supra) and t7SP6BASPOC3 and T7SP6BASPOC4 (Hoegh-Guldberg and Dove, 2000, supra) and fluorescent GFP variants and homologs such as described in Davis and Vierstra, 1996, supra; Haseloff et al., 1997, supra; Lukyanoy et al., 1999, supra; Matz et al., 1999, supra; Fradkov et al., FEBS Letters 479: 127-130, 2000).

- 51 -

Accordingly, another aspect of the present invention provides a eukaryotic or prokaryotic cell or group of eukaryotic or prokaryotic cells genetically modified to comprise:

- (i) a nucleotide sequence set forth in SEQ ID NO:19 or SEQ ID NO:21 or SEQ ID 5 NO:23 or SEQ ID NO:25 or SEQ ID NO:27 or SEQ ID NO:29 or SEQ ID NO:31 or SEQ ID NO:33 or SEQ ID NO:35 or SEQ ID NO:37 or SEQ ID NO:39 or SEQ ID NO:41 or SEQ ID NO:43 or SEQ ID NO:45 or SEQ ID NO:47 or SEQ ID NO:49 or SEQ ID NO:51 or SEQ ID NO:53 or SEQ ID NO:55 or SEQ ID NO:57 or SEQ ID NO:59 or SEQ ID NO:61 or SEQ ID NO:63 or SEQ ID NO:65 or SEQ 10 ID NO:67 or SEQ ID NO:69 or SEQ ID NO:71 or SEQ ID NO:73 or SEQ ID NO:75 or SEQ ID NO:77 or SEQ ID NO:79 or SEQ ID NO:81 or SEQ ID NO:83 or SEQ ID NO:85 or SEQ ID NO:87 or SEQ ID NO:89 or SEQ ID NO:91 or SEQ ID NO:93 or SEQ ID NO:95 or SEQ ID NO:97 or SEQ ID NO:99 or SEQ ID NO:101 or SEQ ID NO:103 or SEQ ID NO:105 or SEQ ID NO:107 or SEQ ID NO:109 or SEQ ID NO:111 or SEQ ID NO:113 or SEQ ID NO:115 or SEQ ID 15 NO:117 or SEQ ID NO:119 or SEQ ID NO:121 or SEQ ID NO:123 or SEQ ID NO:125 or SEQ ID NO:127 or SEQ ID NO:129 or SEQ ID NO:131 or SEQ ID NO:133 or SEQ ID NO:135 or SEQ ID NO:137 or SEQ ID NO:139 or SEQ ID NO:141 or SEQ ID NO:143 or SEQ ID NO:145 or SEQ ID NO:147 or SEQ ID 20 NO:149 or SEQ ID NO:151 or SEQ ID NO:153 or SEQ ID NO:155 or SEQ ID NO:157 or SEQ ID NO:159 or SEQ ID NO:161 or SEQ ID NO:163 or SEQ ID NO:165 or SEQ ID NO:167 or SEQ ID NO:169 or SEQ ID NO:171 or SEQ ID NO:173 or SEQ ID NO:175 or SEQ ID NO:177 or SEQ ID NO:179 or SEQ ID NO:189 or SEQ ID NO:191 or SEQ ID NO:193 or SEQ ID NO:195 or SEQ ID 25 NO:197 or SEQ ID NO:199 or 201;
  - (ii) a nucleotide sequence having at least about 60% similarity after optimal alignment to SEQ ID NO:19 or SEQ ID NO:21 or SEQ ID NO:23 or SEQ ID NO:25 or SEQ ID NO:27 or SEQ ID NO:29 or SEQ ID NO:31 or SEQ ID NO:33 or SEQ ID NO:35 or SEQ ID NO:37 or SEQ ID NO:39 or SEQ ID NO:41 or SEQ ID NO:43 or SEQ ID NO:45 or SEQ ID NO:47 or SEQ ID NO:49 or SEQ ID NO:51 or SEQ

30

- 52 -

5

10

15

ID NO:53 or SEQ ID NO:55 or SEQ ID NO:57 or SEQ ID NO:59 or SEQ ID NO:61 or SEQ ID NO:63 or SEQ ID NO:65 or SEQ ID NO:67 or SEQ ID NO:69 or SEQ ID NO:71 or SEQ ID NO:73 or SEQ ID NO:75 or SEQ ID NO:77 or SEQ ID NO:79 or SEQ ID NO:81 or SEQ ID NO:83 or SEQ ID NO:85 or SEQ ID NO:87 or SEQ ID NO:89 or SEQ ID NO:91 or SEQ ID NO:93 or SEQ ID NO:95 or SEQ ID NO:97 or SEQ ID NO:99 or SEQ ID NO:101 or SEQ ID NO:103 or SEQ ID NO:105 or SEQ ID NO:107 or SEQ ID NO:109 or SEQ ID NO:111 or SEQ ID NO:113 or SEQ ID NO:115 or SEQ ID NO:117 or SEQ ID NO:119 or SEQ ID NO:121 or SEQ ID NO:123 or SEQ ID NO:125 or SEQ ID NO:127 or SEQ ID NO:129 or SEQ ID NO:131 or SEQ ID NO:133 or SEQ ID NO:135 or SEQ ID NO:137 or SEQ ID NO:139 or SEQ ID NO:141 or SEQ ID NO:143 or SEQ ID NO:145 or SEQ ID NO:147 or SEQ ID NO:149 or SEQ ID NO:151 or SEQ ID NO:153 or SEQ ID NO:155 or SEQ ID NO:157 or SEQ ID NO:159 or SEQ ID NO:161 or SEQ ID NO:163 or SEQ ID NO:165 or SEQ ID NO:167 or SEQ ID NO:169 or SEQ ID NO:171 or SEQ ID NO:173 or SEQ ID NO:175 or SEQ ID NO:177 or SEQ ID NO:179 or SEQ ID NO:189 or SEQ ID NO:191 or SEQ ID NO:193 or SEQ ID NO:195 or SEQ ID NO:197 or SEQ ID NO:199 or 201;

20 (iii) a nucleotide sequence capable of hybridizing under low stringency conditions to SEQ ID NO:19 or SEQ ID NO:21 or SEQ ID NO:23 or SEQ ID NO:25 or SEQ ID NO:27 or SEQ ID NO:29 or SEQ ID NO:31 or SEQ ID NO:33 or SEQ ID NO:35 or SEQ ID NO:37 or SEQ ID NO:39 or SEQ ID NO:41 or SEQ ID NO:43 or SEQ ID NO:45 or SEQ ID NO:47 or SEQ ID NO:49 or SEQ ID NO:51 or SEQ ID 25 NO:53 or SEQ ID NO:55 or SEQ ID NO:57 or SEQ ID NO:59 or SEO ID NO:61 or SEQ ID NO:63 or SEQ ID NO:65 or SEQ ID NO:67 or SEQ ID NO:69 or SEQ ID NO:71 or SEQ ID NO:73 or SEQ ID NO:75 or SEQ ID NO:77 or SEQ ID NO:79 or SEQ ID NO:81 or SEQ ID NO:83 or SEQ ID NO:85 or SEQ ID NO:87 or SEQ ID NO:89 or SEQ ID NO:91 or SEQ ID NO:93 or SEQ ID NO:95 or SEQ 30 ID NO:97 or SEQ ID NO:99 or SEQ ID NO:101 or SEQ ID NO:103 or SEQ ID NO:105 or SEQ ID NO:107 or SEQ ID NO:109 or SEQ ID NO:111 or SEQ ID

- 53 -

NO:113 or SEQ ID NO:115 or SEQ ID NO:117 or SEQ ID NO:119 or SEQ ID NO:121 or SEQ ID NO:123 or SEQ ID NO:125 or SEQ ID NO:127 or SEQ ID NO:129 or SEQ ID NO:131 or SEQ ID NO:133 or SEQ ID NO:135 or SEQ ID NO:137 or SEQ ID NO:139 or SEQ ID NO:141 or SEQ ID NO:143 or SEQ ID NO:145 or SEQ ID NO:147 or SEQ ID NO:149 or SEQ ID NO:151 or SEQ ID NO:153 or SEQ ID NO:155 or SEQ ID NO:157 or SEQ ID NO:159 or SEQ ID NO:161 or SEQ ID NO:163 or SEQ ID NO:165 or SEQ ID NO:167 or SEQ ID NO:169 or SEQ ID NO:171 or SEQ ID NO:173 or SEQ ID NO:175 or SEQ ID NO:177 or SEQ ID NO:179 or SEQ ID NO:191 or SEQ ID NO:193 or SEQ ID NO:195 or SEQ ID NO:195 or SEQ ID NO:197 or SEQ ID NO:199 or 201;

5

10

(iv) a nucleotide sequence capable of encoding the amino acid sequence set forth in SEQ ID NO:19 or SEQ ID NO:21 or SEQ ID NO:23 or SEQ ID NO:25 or SEQ ID NO:27 or SEQ ID NO:29 or SEQ ID NO:31 or SEQ ID NO:33 or SEQ ID NO:35 15 or SEQ ID NO:37 or SEQ ID NO:39 or SEQ ID NO:41 or SEQ ID NO:43 or SEQ ID NO:45 or SEQ ID NO:47 or SEQ ID NO:49 or SEQ ID NO:51 or SEQ ID NO:53 or SEQ ID NO:55 or SEQ ID NO:57 or SEQ ID NO:59 or SEQ ID NO:61 or SEQ ID NO:63 or SEQ ID NO:65 or SEQ ID NO:67 or SEQ ID NO:69 or SEQ ID NO:71 or SEQ ID NO:73 or SEQ ID NO:75 or SEQ ID NO:77 or SEQ ID 20 NO:79 or SEQ ID NO:81 or SEQ ID NO:83 or SEQ ID NO:85 or SEQ ID NO:87 or SEQ ID NO:89 or SEQ ID NO:91 or SEQ ID NO:93 or SEQ ID NO:95 or SEQ ID NO:97 or SEQ ID NO:99 or SEQ ID NO:101 or SEQ ID NO:103 or SEQ ID NO:105 or SEQ ID NO:107 or SEQ ID NO:109 or SEQ ID NO:111 or SEQ ID NO:113 or SEQ ID NO:115 or SEQ ID NO:117 or SEQ ID NO:119 or SEQ ID 25 NO:121 or SEQ ID NO:123 or SEQ ID NO:125 or SEQ ID NO:127 or SEQ ID NO:129 or SEQ ID NO:131 or SEQ ID NO:133 or SEQ ID NO:135 or SEQ ID NO:137 or SEQ ID NO:139 or SEQ ID NO:141 or SEQ ID NO:143 or SEQ ID NO:145 or SEQ ID NO:147 or SEQ ID NO:149 or SEQ ID NO:151 or SEQ ID NO:153 or SEQ ID NO:155 or SEQ ID NO:157 or SEQ ID NO:159 or SEQ ID 30 NO:161 or SEQ ID NO:163 or SEQ ID NO:165 or SEQ ID NO:167 or SEQ ID NO:169 or SEQ ID NO:171 or SEQ ID NO:173 or SEQ ID NO:175 or SEQ ID

- 54 -

NO:177 or SEQ ID NO:179 or SEQ ID NO:189 or SEQ ID NO:191 or SEQ ID NO:193 or SEQ ID NO:195 or SEQ ID NO:197 or SEQ ID NO:199 or 201;

- (v) a nucleotide sequence capable of encoding an amino acid sequence having at least 5 about 60% similarity after optimal alignment to SEQ ID NO:19 or SEQ ID NO:21 or SEQ ID NO:23 or SEQ ID NO:25 or SEQ ID NO:27 or SEQ ID NO:29 or SEQ ID NO:31 or SEQ ID NO:33 or SEQ ID NO:35 or SEQ ID NO:37 or SEQ ID NO:39 or SEQ ID NO:41 or SEQ ID NO:43 or SEQ ID NO:45 or SEQ ID NO:47 or SEQ ID NO:49 or SEQ ID NO:51 or SEQ ID NO:53 or SEQ ID NO:55 or SEQ 10 ID NO:57 or SEQ ID NO:59 or SEQ ID NO:61 or SEQ ID NO:63 or SEQ ID NO:65 or SEQ ID NO:67 or SEQ ID NO:69 or SEQ ID NO:71 or SEQ ID NO:73 or SEQ ID NO:75 or SEQ ID NO:77 or SEQ ID NO:79 or SEQ ID NO:81 or SEQ ID NO:83 or SEQ ID NO:85 or SEQ ID NO:87 or SEQ ID NO:89 or SEQ ID NO:91 or SEQ ID NO:93 or SEQ ID NO:95 or SEQ ID NO:97 or SEQ ID NO:99 15 or SEQ ID NO:101 or SEQ ID NO:103 or SEQ ID NO:105 or SEQ ID NO:107 or SEQ ID NO:109 or SEQ ID NO:111 or SEQ ID NO:113 or SEQ ID NO:115 or SEQ ID NO:117 or SEQ ID NO:119 or SEQ ID NO:121 or SEQ ID NO:123 or SEQ ID NO:125 or SEQ ID NO:127 or SEQ ID NO:129 or SEQ ID NO:131 or SEQ ID NO:133 or SEQ ID NO:135 or SEQ ID NO:137 or SEQ ID NO:139 or 20 SEQ ID NO:141 or SEQ ID NO:143 or SEQ ID NO:145 or SEQ ID NO:147 or SEQ ID NO:149 or SEQ ID NO:151 or SEQ ID NO:153 or SEQ ID NO:155 or SEQ ID NO:157 or SEQ ID NO:159 or SEQ ID NO:161 or SEQ ID NO:163 or SEQ ID NO:165 or SEQ ID NO:167 or SEQ ID NO:169 or SEQ ID NO:171 or SEQ ID NO:173 or SEQ ID NO:175 or SEQ ID NO:177 or SEQ ID NO:179 or 25 SEQ ID NO:189 or SEQ ID NO:191 or SEQ ID NO:193 or SEQ ID NO:195 or SEQ ID NO:197 or SEQ ID NO:199 or 201;
  - (vi) a nucleotide sequence capable of hybriding under low stringency conditions to the nucleotide sequence in (iv) or (v) or its complementary form;

- 55 -

wherein said nucleotide sequences encode a CFM which imparts an altered visual characterization to said cell or group of cells to a human eye in the absence of extraneous non-white light or particle emission.

- More particularly, the present invention provides a eukaryotic or prokaryotic cell or group of eukaryotic or prokaryotic cells genetically modified to comprise:
  - (i) a nucleotide sequence set forth in SEQ ID NO:189 or SEQ ID NO:191 or SEQ ID NO:193 or SEQ ID NO:195 or SEQ ID NO:197 or SEQ ID NO:199 or SEQ ID NO:201;
  - (ii) a nucleotide sequence having at least about 60% similarity after optimal alignment to SEQ ID NO:189 or SEQ ID NO:191 or SEQ ID NO:193 or SEQ ID NO:195 or SEQ ID NO:197 or SEQ ID NO:199 or SEQ ID NO:201;

15

10

(iii) a nucleotide sequence capable of hybridizing under low stringency conditions to SEQ ID NO:189 or SEQ ID NO:191 or SEQ ID NO:193 or SEQ ID NO:195 or SEQ ID NO:197 or SEQ ID NO:199 or SEQ ID NO:201 or its complementary form;

20

- (iv) a nucleotide sequence capable of encoding the amino acid sequence set forth in SEQ ID NO:190 or SEQ ID NO:192 or SEQ ID NO:194 or SEQ ID NO:196 or SEQ ID NO:198 or SEQ ID NO:200 or SEQ ID NO:202;
- 25 (v) a nucleotide sequence capable of encoding an amino acid sequence having at least about 60% similarity after optimal alignment to SEQ ID NO:190 or SEQ ID NO:192 or SEQ ID NO:194 or SEQ ID NO:196 or SEQ ID NO:198 or SEQ ID NO:200 or SEQ ID NO:202;
- 30 (vi) a nucleotide sequence capable of hybridizing under low stringency conditions to the nucleotide sequence in (iv) or (v) or its complementary form;

- 56 -

wherein said nucleotide sequences encode a CFM which imparts an altered visual characterization to said cell or group of cells to a human eye in the absence of extraneous non-white light or particle emission.

5

Preferably, the eukaryotic cells are plant cells.

Accordingly, in another aspect of the present invention, there is provided a plant or cells of a plant or parts of a plant or progeny of a plant wherein said plant comprises cells comprising:

(i) a nucleotide sequence set forth in SEQ ID NO:19 or SEQ ID NO:21 or SEQ ID NO:23 or SEQ ID NO:25 or SEQ ID NO:27 or SEQ ID NO:29 or SEQ ID NO:31 or SEQ ID NO:33 or SEQ ID NO:35 or SEQ ID NO:37 or SEQ ID NO:39 or SEQ 15 ID NO:41 or SEQ ID NO:43 or SEQ ID NO:45 or SEQ ID NO:47 or SEQ ID NO:49 or SEQ ID NO:51 or SEQ ID NO:53 or SEQ ID NO:55 or SEQ ID NO:57 or SEQ ID NO:59 or SEQ ID NO:61 or SEQ ID NO:63 or SEQ ID NO:65 or SEQ ID NO:67 or SEQ ID NO:69 or SEQ ID NO:71 or SEQ ID NO:73 or SEQ ID NO:75 or SEQ ID NO:77 or SEQ ID NO:79 or SEQ ID NO:81 or SEQ ID NO:83 20 or SEQ ID NO:85 or SEQ ID NO:87 or SEQ ID NO:89 or SEQ ID NO:91 or SEQ ID NO:93 or SEQ ID NO:95 or SEQ ID NO:97 or SEQ ID NO:99 or SEQ ID NO:101 or SEQ ID NO:103 or SEQ ID NO:105 or SEQ ID NO:107 or SEQ ID NO:109 or SEQ ID NO:111 or SEQ ID NO:113 or SEQ ID NO:115 or SEQ ID NO:117 or SEQ ID NO:119 or SEQ ID NO:121 or SEQ ID NO:123 or SEQ ID 25 NO:125 or SEQ ID NO:127 or SEQ ID NO:129 or SEQ ID NO:131 or SEQ ID NO:133 or SEQ ID NO:135 or SEQ ID NO:137 or SEQ ID NO:139 or SEQ ID NO:141 or SEQ ID NO:143 or SEQ ID NO:145 or SEQ ID NO:147 or SEQ ID NO:149 or SEQ ID NO:151 or SEQ ID NO:153 or SEQ ID NO:155 or SEQ ID NO:157 or SEQ ID NO:159 or SEQ ID NO:161 or SEQ ID NO:163 or SEQ ID 30 NO:165 or SEQ ID NO:167 or SEQ ID NO:169 or SEQ ID NO:171 or SEQ ID NO:173 or SEQ ID NO:175 or SEQ ID NO:177 or SEQ ID NO:179 or SEQ ID

- 57 -

NO:189 or SEQ ID NO:191 or SEQ ID NO:193 or SEQ ID NO:195 or SEQ ID NO:197 or SEQ ID NO:199 or 201;

- (ii) a nucleotide sequence having at least about 60% similarity after optimal alignment 5 to SEQ ID NO:19 or SEQ ID NO:21 or SEQ ID NO:23 or SEQ ID NO:25 or SEQ ID NO:27 or SEQ ID NO:29 or SEQ ID NO:31 or SEQ ID NO:33 or SEQ ID NO:35 or SEQ ID NO:37 or SEQ ID NO:39 or SEQ ID NO:41 or SEQ ID NO:43 or SEQ ID NO:45 or SEQ ID NO:47 or SEQ ID NO:49 or SEQ ID NO:51 or SEQ ID NO:53 or SEQ ID NO:55 or SEQ ID NO:57 or SEQ ID NO:59 or SEQ ID 10 NO:61 or SEQ ID NO:63 or SEQ ID NO:65 or SEQ ID NO:67 or SEQ ID NO:69 or SEQ ID NO:71 or SEQ ID NO:73 or SEQ ID NO:75 or SEQ ID NO:77 or SEQ ID NO:79 or SEQ ID NO:81 or SEQ ID NO:83 or SEQ ID NO:85 or SEQ ID NO:87 or SEQ ID NO:89 or SEQ ID NO:91 or SEQ ID NO:93 or SEQ ID NO:95 or SEQ ID NO:97 or SEQ ID NO:99 or SEQ ID NO:101 or SEQ ID NO:103 or 15 SEQ ID NO:105 or SEQ ID NO:107 or SEQ ID NO:109 or SEQ ID NO:111 or SEQ ID NO:113 or SEQ ID NO:115 or SEQ ID NO:117 or SEQ ID NO:119 or SEQ ID NO:121 or SEQ ID NO:123 or SEQ ID NO:125 or SEQ ID NO:127 or SEQ ID NO:129 or SEQ ID NO:131 or SEQ ID NO:133 or SEQ ID NO:135 or SEQ ID NO:137 or SEQ ID NO:139 or SEQ ID NO:141 or SEQ ID NO:143 or 20 SEQ ID NO:145 or SEQ ID NO:147 or SEQ ID NO:149 or SEQ ID NO:151 or SEQ ID NO:153 or SEQ ID NO:155 or SEQ ID NO:157 or SEQ ID NO:159 or SEQ ID NO:161 or SEQ ID NO:163 or SEQ ID NO:165 or SEQ ID NO:167 or SEQ ID NO:169 or SEQ ID NO:171 or SEQ ID NO:173 or SEQ ID NO:175 or SEQ ID NO:177 or SEQ ID NO:179 or SEQ ID NO:189 or SEQ ID NO:191 or 25 SEQ ID NO:193 or SEQ ID NO:195 or SEQ ID NO:197 or SEQ ID NO:199 or 201;
- (iii) a nucleotide sequence capable of hybridizing under low stringency conditions to SEQ ID NO:19 or SEQ ID NO:21 or SEQ ID NO:23 or SEQ ID NO:25 or SEQ ID NO:35 or SEQ ID NO:29 or SEQ ID NO:31 or SEQ ID NO:33 or SEQ ID NO:35 or SEQ ID NO:37 or SEQ ID NO:39 or SEQ ID NO:41 or SEQ ID NO:43 or SEQ

- 58 -

5

10

15

ID NO:45 or SEQ ID NO:47 or SEQ ID NO:49 or SEQ ID NO:51 or SEQ ID NO:53 or SEQ ID NO:55 or SEQ ID NO:57 or SEQ ID NO:59 or SEQ ID NO:61 or SEQ ID NO:63 or SEQ ID NO:65 or SEQ ID NO:67 or SEQ ID NO:69 or SEQ ID NO:71 or SEQ ID NO:73 or SEQ ID NO:75 or SEQ ID NO:77 or SEQ ID NO:79 or SEQ ID NO:81 or SEQ ID NO:83 or SEQ ID NO:85 or SEQ ID NO:87 or SEQ ID NO:89 or SEQ ID NO:91 or SEQ ID NO:93 or SEQ ID NO:95 or SEQ ID NO:97 or SEQ ID NO:99 or SEQ ID NO:101 or SEQ ID NO:103 or SEQ ID NO:105 or SEQ ID NO:107 or SEQ ID NO:109 or SEQ ID NO:111 or SEO ID NO:113 or SEQ ID NO:115 or SEQ ID NO:117 or SEQ ID NO:119 or SEQ ID NO:121 or SEQ ID NO:123 or SEQ ID NO:125 or SEQ ID NO:127 or SEQ ID NO:129 or SEQ ID NO:131 or SEQ ID NO:133 or SEQ ID NO:135 or SEQ ID NO:137 or SEQ ID NO:139 or SEQ ID NO:141 or SEQ ID NO:143 or SEQ ID NO:145 or SEQ ID NO:147 or SEQ ID NO:149 or SEQ ID NO:151 or SEQ ID NO:153 or SEQ ID NO:155 or SEQ ID NO:157 or SEQ ID NO:159 or SEQ ID NO:161 or SEQ ID NO:163 or SEQ ID NO:165 or SEQ ID NO:167 or SEQ ID NO:169 or SEQ ID NO:171 or SEQ ID NO:173 or SEQ ID NO:175 or SEQ ID NO:177 or SEQ ID NO:179 or SEQ ID NO:189 or SEQ ID NO:191 or SEQ ID NO:193 or SEQ ID NO:195 or SEQ ID NO:197 or SEQ ID NO:199 or 201;

20 (iv) a nucleotide sequence capable of encoding the amino acid sequence set forth in SEQ ID NO:19 or SEQ ID NO:21 or SEQ ID NO:23 or SEQ ID NO:25 or SEQ ID NO:27 or SEQ ID NO:29 or SEQ ID NO:31 or SEQ ID NO:33 or SEQ ID NO:35 or SEQ ID NO:37 or SEQ ID NO:39 or SEQ ID NO:41 or SEQ ID NO:43 or SEQ ID NO:45 or SEQ ID NO:47 or SEQ ID NO:49 or SEQ ID NO:51 or SEQ ID 25 NO:53 or SEQ ID NO:55 or SEQ ID NO:57 or SEQ ID NO:59 or SEQ ID NO:61 or SEQ ID NO:63 or SEQ ID NO:65 or SEQ ID NO:67 or SEQ ID NO:69 or SEQ ID NO:71 or SEQ ID NO:73 or SEQ ID NO:75 or SEQ ID NO:77 or SEQ ID NO:79 or SEQ ID NO:81 or SEQ ID NO:83 or SEQ ID NO:85 or SEQ ID NO:87 or SEQ ID NO:89 or SEQ ID NO:91 or SEQ ID NO:93 or SEQ ID NO:95 or SEQ 30 ID NO:97 or SEQ ID NO:99 or SEQ ID NO:101 or SEQ ID NO:103 or SEQ ID NO:105 or SEQ ID NO:107 or SEQ ID NO:109 or SEQ ID NO:111 or SEQ ID

- 59 -

NO:113 or SEQ ID NO:115 or SEQ ID NO:117 or SEQ ID NO:119 or SEQ ID NO:121 or SEQ ID NO:123 or SEQ ID NO:125 or SEQ ID NO:127 or SEQ ID NO:127 or SEQ ID NO:129 or SEQ ID NO:131 or SEQ ID NO:133 or SEQ ID NO:135 or SEQ ID NO:135 or SEQ ID NO:137 or SEQ ID NO:139 or SEQ ID NO:141 or SEQ ID NO:143 or SEQ ID NO:145 or SEQ ID NO:147 or SEQ ID NO:149 or SEQ ID NO:151 or SEQ ID NO:153 or SEQ ID NO:155 or SEQ ID NO:157 or SEQ ID NO:159 or SEQ ID NO:161 or SEQ ID NO:163 or SEQ ID NO:165 or SEQ ID NO:165 or SEQ ID NO:167 or SEQ ID NO:169 or SEQ ID NO:171 or SEQ ID NO:173 or SEQ ID NO:175 or SEQ ID NO:177 or SEQ ID NO:179 or SEQ ID NO:191 or SEQ ID NO:193 or SEQ ID NO:195 or SEQ ID NO:195 or SEQ ID NO:197 or SEQ ID NO:199 or 201;

5

10

15

20

25

30

a nucleotide sequence capable of encoding an amino acid sequence having at least (v) about 60% similarity after optimal alignment to SEQ ID NO:19 or SEQ ID NO:21 or SEQ ID NO:23 or SEQ ID NO:25 or SEQ ID NO:27 or SEQ ID NO:29 or SEO ID NO:31 or SEQ ID NO:33 or SEQ ID NO:35 or SEQ ID NO:37 or SEQ ID NO:39 or SEQ ID NO:41 or SEQ ID NO:43 or SEQ ID NO:45 or SEQ ID NO:47 or SEQ ID NO:49 or SEQ ID NO:51 or SEQ ID NO:53 or SEQ ID NO:55 or SEQ ID NO:57 or SEQ ID NO:59 or SEQ ID NO:61 or SEQ ID NO:63 or SEQ ID NO:65 or SEQ ID NO:67 or SEQ ID NO:69 or SEQ ID NO:71 or SEQ ID NO:73 or SEQ ID NO:75 or SEQ ID NO:77 or SEQ ID NO:79 or SEQ ID NO:81 or SEQ ID NO:83 or SEQ ID NO:85 or SEQ ID NO:87 or SEQ ID NO:89 or SEQ ID NO:91 or SEQ ID NO:93 or SEQ ID NO:95 or SEQ ID NO:97 or SEQ ID NO:99 or SEQ ID NO:101 or SEQ ID NO:103 or SEQ ID NO:105 or SEQ ID NO:107 or SEQ ID NO:109 or SEQ ID NO:111 or SEQ ID NO:113 or SEQ ID NO:115 or SEQ ID NO:117 or SEQ ID NO:119 or SEQ ID NO:121 or SEQ ID NO:123 or SEQ ID NO:125 or SEQ ID NO:127 or SEQ ID NO:129 or SEQ ID NO:131 or SEQ ID NO:133 or SEQ ID NO:135 or SEQ ID NO:137 or SEQ ID NO:139 or SEQ ID NO:141 or SEQ ID NO:143 or SEQ ID NO:145 or SEQ ID NO:147 or SEQ ID NO:149 or SEQ ID NO:151 or SEQ ID NO:153 or SEQ ID NO:155 or SEQ ID NO:157 or SEQ ID NO:159 or SEQ ID NO:161 or SEQ ID NO:163 or SEQ ID NO:165 or SEQ ID NO:167 or SEQ ID NO:169 or SEQ ID NO:171 or

- 60 -

SEQ ID NO:173 or SEQ ID NO:175 or SEQ ID NO:177 or SEQ ID NO:179 or SEQ ID NO:189 or SEQ ID NO:191 or SEQ ID NO:193 or SEQ ID NO:195 or SEQ ID NO:197 or SEQ ID NO:199 or 201;

5 (vi) a nucleotide sequence capable of hybriding under low stringency conditions to the nucleotide sequence in (iv) or (v) or its complementary form;

wherein said nucleotide sequences encode a CFM which imparts an altered visual characterization to said cell or group of cells to a human eye in the absence of extraneous non-white light or particle emission.

10

25

More particularly, there is provided a plant or cells of a plant or parts of a plant or progeny of a plant wherein said plant comprises cells comprising:

- 15 (i) a nucleotide sequence set forth in SEQ ID NO:189 or SEQ ID NO:191 or SEQ ID NO:193 or SEQ ID NO:195 or SEQ ID NO:197 or SEQ ID NO:199 or SEQ ID NO:201;
- (ii) a nucleotide sequence having at least about 60% similarity after optimal alignment to SEQ ID NO:189 or SEQ ID NO:191 or SEQ ID NO:193 or SEQ ID NO:195 or SEQ ID NO:197 or SEQ ID NO:199 or SEQ ID NO:201;
  - (iii) a nucleotide sequence capable of hybridizing under low stringency conditions to SEQ ID NO:189 or SEQ ID NO:191 or SEQ ID NO:193 or SEQ ID NO:195 or SEQ ID NO:197 or SEQ ID NO:199 or SEQ ID NO:201 or its complementary form;
- (iv) a nucleotide sequence capable of encoding the amino acid sequence set forth in SEQ ID NO:190 or SEQ ID NO:192 or SEQ ID NO:194 or SEQ ID NO:196 or SEQ ID NO:198 or SEQ ID NO:200 or SEQ ID NO:202;

- 61 -

(v) a nucleotide sequence capable of encoding an amino acid sequence having at least about 60% similarity after optimal alignment SEQ ID NO:190 or SEQ ID NO:192 or SEQ ID NO:194 or SEQ ID NO:196 or SEQ ID NO:198 or SEQ ID NO:200 or SEQ ID NO:202;

5

20

30

(vi) a nucleotide sequence capable of hybridizing under low stringency conditions to the nucleotide sequence in (iv) or (v) or its complementary form;

wherein said nucleotide sequences encode a CFM which imparts an altered visual 10 characterization to said plant or cells of a plant to a human eye in the absence of extraneous non-white light or particle emission.

In a particularly preferred embodiment, there is provided a use of a CFM such as but not limited to GFP or a non-fluorescent GFP-homolog in the manufacture of a plant exhibiting altered visual characteristics to all or a part of said plant or to cells of said plant to a human eye in the absence of extraneous non-white light or particle emission.

Reference herein to extraneous light is not to be read as encompassing white light or background irradiation. The altered visual characteristics are visualized in the presence of white light, for example the light as generated by an 60 W electric bulb or daylight. White light includes light that contains all the wavelengths of the visible spectrum, such as sunlight.

The term "similarity" as used herein includes exact identity between compared sequences 25 at the nucleotide or amino acid level. Where there is non-identity at the nucleotide level, "similarity" includes differences between sequences which result in different amino acids that are nevertheless related to each other at the structural, functional, biochemical and/or conformational levels. Where there is non-identity at the amino acid level, "similarity" includes amino acids that are nevertheless related to each other at the structural, functional, biochemical and/or conformational levels. In a particularly preferred embodiment,

- 62 -

nucleotide and sequence comparisons are made at the level of identity rather than similarity.

Terms used to describe sequence relationships between two or more polynucleotides or polypeptides include "reference sequence", "comparison window", "sequence similarity", "sequence identity", "percentage of sequence similarity", "percentage of sequence identity", "substantially similar" and "substantial identity". A "reference sequence" is at least 12 but frequently 15 to 18 and often at least 25 or above, such as 30 monomer units, inclusive of nucleotides and amino acid residues, in length. Because two polynucleotides may each comprise (1) a sequence (i.e. only a portion of the complete polynucleotide sequence) that is similar between the two polynucleotides, and (2) a sequence that is divergent between the two polynucleotides, sequence comparisons between two (or more) polynucleotides are typically performed by comparing sequences of the two polynucleotides over a "comparison window" to identify and compare local regions of sequence similarity. A "comparison window" refers to a conceptual segment of typically 12 contiguous residues that is compared to a reference sequence. The comparison window may comprise additions or deletions (i.e. gaps) of about 20% or less as compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. Optimal alignment of sequences for aligning a comparison window may be conducted by computerized implementations of algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package Release 7.0, Genetics Computer Group, 575 Science Drive Madison, WI, USA) or by inspection and the best alignment (i.e. resulting in the highest percentage homology over the comparison window) generated by any of the various methods selected. Reference also may be made to the BLAST family of programs as for example disclosed by Altschul et al. (Nucl. Acids Res. 25: 3389, 1997). A detailed discussion of sequence analysis can be found in Unit 19.3 of Ausubel et al. (Current Protocols in Molecular Biology, John Wiley & Sons Inc, 1994-1998, Chapter 15).

10

15

30 The terms "sequence similarity" and "sequence identity" as used herein refers to the extent that sequences are identical or functionally or structurally similar on a nucleotide-by-

- 63 -

nucleotide basis or an amino acid-by-amino acid basis over a window of comparison. Thus, a "percentage of sequence identity", for example, is calculated by comparing two optimally aligned sequences over the window of comparison, determining the number of positions at which the identical nucleic acid base (e.g. A, T, C, G, I) or the identical amino acid residue (e.g. Ala, Pro, Ser, Thr, Gly, Val, Leu, Ile, Phe, Tyr, Trp, Lys, Arg, His, Asp, Glu, Asn, Gln, Cys and Met) occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison (i.e., the window size), and multiplying the result by 100 to yield the percentage of sequence identity. For the purposes of the present invention, "sequence identity" will be understood to mean the "match percentage" calculated by the DNASIS computer program (Version 2.5 for windows; available from Hitachi Software engineering Co., Ltd., South San Francisco, California, USA) using standard defaults as used in the reference manual accompanying the software. Similar comments apply in relation to sequence similarity.

15

20

10

Reference herein to a low stringency includes and encompasses from at least about 0 to at least about 15% v/v formamide and from at least about 1 M to at least about 2 M salt for hybridization, and at least about 1 M to at least about 2 M salt for washing conditions. Generally, low stringency is at from about 25-30°C to about 42°C. The temperature may be altered and higher temperatures used to replace formamide and/or to give alternative stringency conditions. Alternative stringency conditions may be applied where necessary, such as medium stringency, which includes and encompasses from at least about 16% y/y to at least about 30% v/v formamide and from at least about 0.5 M to at least about 0.9 M salt for hybridization, and at least about 0.5 M to at least about 0.9 M salt for washing conditions, or high stringency, which includes and encompasses from at least about 31% v/v to at least about 50% v/v formamide and from at least about 0.01 M to at least about 0.15 M salt for hybridization, and at least about 0.01 M to at least about 0.15 M salt for washing conditions. In general, washing is carried out  $T_m = 69.3 + 0.41$  (G+C)% (Marmur and Doty, J. Mol. Biol. 5: 109, 1962). However, the T<sub>m</sub> of a duplex DNA decreases by 1°C with every increase of 1% in the number of mismatch base pairs (Bonner and Laskey, Eur. J. Biochem. 46: 83, 1974). Formaride is optional in these hybridization conditions.

- 64 -

Accordingly, particularly preferred levels of stringency are defined as follows: low stringency is 6 x SSC buffer, 0.1% w/v SDS at 25°-42°C; a moderate stringency is 2 x SSC buffer, 0.1% w/v SDS at a temperature in the range 20°C to 65°C; high stringency is 0.1 x SSC buffer, 0.1% w/v SDS at a temperature of at least 65°C.

5

10

The tobacco ribosomal DNA spacer element may be used to increase the expression of CFMs or colored proteins in transgenic *Arabidopsis*, carnation, rose or other plant species. The tobacco ribosomal DNA spacer element can be used to increase copy number and expression levels of transgenes in plants (Borisjuk *et al.*, *Nat. Biotechnol. 18:* 1303-1306, 2000). The tobacco ribosomal DNA spacer element may be inserted into pCGP2772, pCGP2785, pCGP3259 or other construct used to express CFMs or colored proteins in plants.

There is a clear correlation between codon usage and gene expression levels in Arabidopsis, Caenorhabditis and Drosophila (Duret and Mouchiroud, Proc. Natl. Acad. Sci. USA 96: 4482-4487, 1999).

Codon usage within the open reading frames of CFM or colored proteins may be modified to increase levels of CFMs or colored protein in transgenic *Arabidopsis*, carnation, rose or other plant species.

A recent study by Stevens et al. (Plant Physiology 173-182, 2000) has highlighted the possibility of increasing the stability of recombinant proteins in transgenic plants by modifying protein glycosylation patterns.

25

20

Plant virus gene vectors may be used for high level gene expression of foreign genes in plants (Scholthof and Scholthof, Annu. Rev. of Phytopathol. 34: 299-323, 1996; Chapman et al., Plant Journal 2: 549-557, 1992).

30 The use of a plant virus expression system may increase levels of CFMs or colored protein in transgenic *Arabidopsis*, carnation, rose or other plant species. Selection of an

- 65 -

appropriate virus type or strain may allow the expression of CFMs or colored protein in specific tissues or patterns to produce novel phenotypes. For example a CFM or colored protein gene may be incorporated into the genome of tulip breaking virus or tulip chlorotic blotch potyvirus to induce colored sector production in tulip or other flowers.

5

The availability of the isolated CFMs of the present invention further provides the possibility for generating antibodies, whether monoclonal or polyclonal, against any or all of these isolated sequences or derivatives or homologs thereof.

Well-known protocols applicable to antibody production, purification and use may be found, for example, in Chapter 2 of Coligan et al. (Current Protocols in Immunology, John Wiley & Sons, N.Y., 1991-94) and Harlow and Lane, Antibodies: A Laboratory Manual, Cold Spring Harbor, Cold Spring Harbor Laboratory, 1988, which are both herein incorporated by reference.

15

20

Generally, antibodies of the invention bind to or conjugate with a polypeptide, fragment, variant or derivative thereof. For example, the antibodies may comprise polyclonal antibodies. Such antibodies may be prepared, for example, by injecting a polypeptide, fragment, variant or derivative thereof into a production species, which may include mice or rabbits, to obtain polyclonal antisera. Methods for the production of polyclonal antibodies are well known to those skilled in the art. Exemplary protocols are described in Coligan et al., 1991-1994, supra and Harlow and Lane, 1988, supra.

25 be Jour

30

In lieu of polyclonal antisera obtained in a production species, monoclonal antibodies may be produced using the standard method as described by Köhler & Milstein (European Journal of Immunology 6: 511-519, 1976) or by more recent modifications thereof as, for example, described in Coligan et al. (1991-1994, supra) by immortalizing spleen or other antibody-producing cells derived from a production species which has been inoculated with one or more of the polypeptides, fragments, variants or derivatives of the present invention.

- 66 -

The present invention also contemplates antibodies that comprise Fc or Fab fragments of the polyclonal or monoclonal antibodies referred to above. Alternatively, the antibodies may comprise single chain Fv antibodies (scFvs) against the peptides of the present invention. Such scFvs may be prepared, for example, in accordance with the methods described respectively in U.S. Patent No. 5,091,513, European Patent No 239,400 or Winter and Milstein (*Nature 349*: 293, 1991).

Antibodies produced in accordance with the present invention may be used for affinity chromatography in isolating natural or recombinant pigment polypeptides. For appropriate protocols, reference may be made to immunoaffinity chromatographic procedures described in Chapter 9.5 of Coligan et al. (1991-1994, supra).

10

15

Accordingly, the present invention provides an antibody specific for a CFM, said CFM comprising an amino acid sequence in its N-terminal end selected from SVIAK (SEQ ID NO:5), (M)SVIAT (SEQ ID NO:6), SGIAT (SEQ ID NO:7), SVIVT (SEQ ID NO:8) or SVSAT (SEQ ID NO:9).

Preferably, the isolated antibody is specific for a CFM comprising an amino acid sequence selected from the list comprising SVIAT QMTY KVYM SGT (SEQ ID NO:10), SVIAT QMTY KVYM PGT (SEQ ID NO:11), SVIAT QVTY KVYM SGT (SEQ ID NO:12), SGIAT QMTY KVYM SGT (SEQ ID NO:13), SVIVT QMTY KVYM SGT (SEQ ID NO:14), SVSAT QMTY KVYM SGT (SEQ ID NO:15), SVIAK QMTY KVNM SGT (SEQ ID NO:16), SVIAK QMTY KVYM SDT (SEQ ID NO:17) and/or SVIAK QMTY X<sub>1</sub>X<sub>2</sub>YX<sub>3</sub> SGT (SEQ ID NO:18) wherein X<sub>1</sub>, X<sub>2</sub> and X<sub>3</sub> may be any amino acid provided that X<sub>1</sub> is not K; X<sub>2</sub> is not V; X<sub>3</sub> is not M.

Most preferably, the antibody is specific for a CFM comprising an amino acid sequence selected from the listing comprising SEQ ID NOs:20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126,

- 67 -

128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178, 180, 190, 192, 194, 196, 198, 200 and 202.

Once antibodies have been produced, one or more polypeptides of the present invention may be conjugated thereto, preferably to a secondary antibody as part of an antibody staining complex, and thereby become useful as a fluorescent marker in microscopy and related procedures. Alternatively, or in addition, one or more nucleic acid sequence encoding a polypeptide of the present invention may be expressed as a recombinant polypeptide fused with a secondary antibody. These antibodies may be useful for *in situ* labelling procedures or in other related procedures such as fluorescence *in situ* hybridization (FISH).

10

15

As already described above, a fusion partner well known in the art is GFP. This fusion partner may serve as a fluorescent "tag" which facilitates the identification and/or localization, by fluorescence microscopy or by flow cytometry, of a polypeptide fused thereto. Flow cytometric methods such as fluorescence activated cell sorting (FACS) are particularly useful in this regard.

There is perpetual interest in developing high-sensitivity biochemical assays, which 20 employ luminescence, fluorescence or visible color rather than radioisotopes, for use in research and in medicine. Interest in developing assays with visible detection systems is increasing as these often obviate the need for expensive luminescence, fluorescence or isotopic detection equipment.

Accordingly, the present invention further comprises a diagnostic assay comprising screening for the presence of CFM wherein the nucleic acid molecule encoding said CFM is expressed in a cell.

The capability of the CFMs to absorb incident light which encompasses the UV range (320-700 nm) makes them useful candidates for inclusion as components in topically-applicable sun screen formulations. The purpose of a sun screen is to block the excessive

- 68 -

UV radiation from affecting the skin. Sun screen formulations act by deflecting and scattering the incident light that produces burning and tanning of the skin or by absorbing this light. It is known that careful selection of sun screens can offer this protection to the skin and reduce the darkening and damaging effects of the radiation.

5

Such a formulation would include, for example, an effective amount of one or more CFMs of the present invention, optionally admixed with a pharmaceutically acceptable vehicle such as a carrier or excipient that will not harm the skin. By "carrier" is meant a solid or liquid filler, diluent or substance that may be safely used in topical administration. These carriers may be selected from a group including powder absorbants, creams, oils, synthetic oils, phosphate buffered solutions, emulsifiers, and liquids such as emollients, propellants, solvents, humectants, thickners, isotonic saline, and pyrogen-free water. The sun screen formulation may also include other screening agents, well known in the art, such as propyl hydroxybenzoate, dimethylaminobenzoate (PABA), phenyl salicylates and/or octyl methoxycinnamate. These formulations may be prepared for topical application to the skin in the form of conventional products such as lotions, creams, ointments and aerosol products. A useful sun screen formulation and method of preparing an emulsion therefor are provided in International Patent Publication No. WO 00/46233 in Example 4.

- Accordingly, the present invention provides a biomatrix comprising a CFM, said CFM comprising a polypeptide which, in a cell, alone or together with one or more other molecules imparts an altered visual characteristic to said cell when visualized by a human eye in the absence of excitation by extraneous non-white light or particle emission.
- Reference to a "biomatrix" includes any composition comprising a CFM such as a cell, sun screen, a purified preparation of a CFM or any solid support onto or into which a CFM is immobilized. Reference to a biomatrix also includes a bioinstrument.

Yet another aspect of the present invention contemplates the use of a CFM in a cosmetic or light filtering composition. Cosmetics include many products that can be applied to the face or body in order to alter appearance or color. New combinations of ingredients may

- 69 -

result in cosmetic compositions that protect against environmental stresses such as exposure to the sun. The use of a CFM in a cosmetic may provide a visible coloration that is aesthetically desirable and/or it may provide light filtering capability such as may be afforded, for example, by a sun screen.

5

Light filtering compositions may also be used to screen out or block UV light or different wavelengths of light within the entire spectrum. A cosmetic or light filtering composition according to the invention may also include cosmetically or pharmaceutically compatible carriers, preservatives, emusifiers, thickners, perfume, color, as well as other materials having properties which are beneficial for skin, such as moisturizers, emollients antiageing compounds inter alia.

Other applications of the CFMs of the present invention may also be contemplated. Since they are active in affecting the manner in which, and degree to which, various kinds of impinging light/radiation are processed and detected, the CFMs may find application in, for example, transducing or intensifying an image. For example, converting less visible wavelengths of light such as UV radiation to wavelengths that are more visible might be beneficial. A gel or similar material comprising a CFM may be located behind a membrane or selective barrier and combined with an optic fiber probe, such as an optode or microelectrode. Changes in physical and chemical environments into which the probe is inserted may be calibrated to changes in fluorescent intensity and/or fluorescence half-life, to provide micro-scale measurements of parameters such as oxygen concentration and pH. Similar applications involving fluorescence intensity and/or half-life fluorescent imaging techniques may also incorporate a CFM of the present invention.

25

30

15

20

As stated above, each of the CFMs of the present invention and homologs thereof, has distinct excitation and emission characteristics. These may be fluorescently coupled such that captured photons can be passed successively between a plurality of CFMs, for example as many as six. This lengthens the pathway and the amount of time that a photon spends within any material comprising the CFMs and may thereby increase light intensity within these environments considerably. Such a light enhancement effect may be useful for

- 70 -

providing additional light for growing phototrophic organisms, for example plants, algae and/or corals, by increasing the likelihood of a photon's interaction with constituent photosystems.

This embodiment of the present invention may also be useful for creating light enhancer fluids that may be used to increase light intensity within a medium above that of incident light.

Furthermore, a CFM embedded in a gel matrix or other useful material may improve image quality in situations of distorted light spectra such as, for example, under water where light is shifted to the blue end of the spectrum. A CFM rendered water-soluble may prove useful in a range of different types of liquids. Alternatively, or in addition, a derivative or homolog of polypeptide of the present invention may be synthesised by substituting amino acids or adding N- or C-terminal tags to increase their insolubility and hence make them more useful in less polar environments. In this embodiment, a CFM, or a CFM modified such as through amino acid inclusion or substitution to make it more hydrophobic, combined with a water-soluble or non-water soluble emulsion, may be used to coat materials that experience UV damage such as, for example, plastics and car upholstery.

20 The present invention is further described by the following non-limiting Examples.

10

15

- 71 -

## **EXAMPLE 1**

## General methods

In general, the methods followed were as described in Sambrook et al. (Molecular Cloning: A Laboratory Manual. (2nd edition), Cold Spring Harbor Laboratory Press, USA, 1989).

The cloning vectors pBluescript and PCR script were obtained from Stratagene. pCR7 2.1 was obtained from Invitrogen.

10

The bacterial expression vector pQE-30 was obtained from Qiagen.

## E. coli transformation

5 The Escherichia coli strains used were:-

DH5 $\alpha$ 

supE44,  $\triangle$  (lacZYA-ArgF)U169, (ø80lacZ $\triangle$ M15), hsdR17( $r_k$ ,  $m_k$ <sup>+</sup>), recA1, endA1, gyrA96, thi-1, relA1, deoR. (Hanahan, J. Mol. Biol. 166: 557 1983

:0

Ĩ

2 XL1-Blue

supE44, hsdR17(r<sub>k</sub>, m<sub>k</sub>), recA1, endA1, gyrA96, thi-1, relA1, lac, [FproAB, lacI<sup>q</sup>, lacZΔM15, Tn10(tet<sup>R</sup>)] (Bullock et al., Biotechniques 5: 376, 1987).

- 5 BL21-CodonPlus-RIL strain
- ompT hsdS(rB-mB-) dcm+ Tet<sup>r</sup> gal endA Hte [argU ileY leuW Cam<sup>r</sup>]

  M15 E. coli is derived from E.coli K12 and has the phenotype Nal<sup>s</sup>, Str<sup>s</sup>, Rif<sup>s</sup>, Thi, Ara<sup>+</sup>, Gal<sup>+</sup>, Mtl, F, RecA<sup>+</sup>, Uvr<sup>+</sup>, Lon<sup>+</sup>.
- O Transformation of the *E. coli* strains was performed according to the method of Inoue *et al.*, (Gene 96: 23-28, 1990).

- 72 -

#### Agrobacterium tumefaciens strains and transformations

The disarmed Agrobacterium tumefaciens strain used was AGL0 (Lazo et al. Bio/technology 9: 963-967, 1991).

Plasmid DNA was introduced into the Agrobacterium tumefaciens strain AGL0 by adding 5 µg of plasmid DNA to 100 µL of competent AGL0 cells prepared by inoculating a 50 mL LB culture and growing for 16 hours with shaking at 28°C. The cells were then pelleted and resuspended in 0.5mL of 85% v/v 100mM CaCl<sub>2</sub>/15% v/v) glycerol. The DNA-Agrobacterium mixture was frozen by incubation in liquid N<sub>2</sub> for 2 minutes and then allowed to thaw by incubation at 37°C for 5 minutes. The DNA/bacterial mix was then placed on ice for a further 10 minutes. The cells were then mixed with 1mL of LB (Sambrook et al., 1989 supra) media and incubated with shaking for 16 hours at 28°C. Cells of A. tumefaciens carrying the plasmid were selected on LB agar plates containing 50 µg/mL tetracycline. The confirmation of the plasmid in A. tumefaciens was done by restriction enzyme analysis of DNA isolated from the tetracycline-resistant transformants.

# Saccharomyces cerevisiae strains and transformations

20

5

10

15

The yeast expression vector used was pYE22m (Tanaka et al., J. Biochem. 103: 954-961, 1988).

The yeast strain G-1315 (Mat α trpl) (Ashikari et al., Appl. Microbiol. Biotechnol. 30: 515-520, 1989) was transformed with plasmid DNA according to Ito et al., (J. Bacteriol. 153: 163-168, 1983). The transformants were selected by their ability to restore G-1315 to tryptophan prototrophy.

- 73 -

#### DNA ligations

DNA ligations were carried out using the Amersham Ligation Kit according to procedures recommended by the manufacturer.

5

#### Isolation and purification of DNA fragments

Fragments were generally isolated on a 1% w/v agarose gel and purified using the QIAEX II Gel Extraction kit (Qiagen).

10

15

20

## Reparation of overhanging ends after restriction digestion

Overhanging 5' ends were repaired using DNA polymerase (Klenow fragment) according to standard protocols (Sambrook et al., 1989 supra). Overhanging 3' ends were repaired using T4 DNA polymerase according to standard protocols (Sambrook et al., 1989 supra).

## Removal of phosphoryl groups from nucleic acids

Shrimp alkaline phosphatase (SAP) (USB) was typically used to remove phosphoryl groups from cloning vectors to prevent re-circularization according to the manufacturer's recommendations.

#### Polymerase Chain Reaction (PCR)

Unless otherwise specified, PCR conditions using plasmid DNA as template included using 2ng plasmid, 100ng each of primers, 2 μL 10 mM dNTP mix, 5 μL 10 x PfuTurbo (registered trademark) DNA polymerase buffer (Stratageme), 0.5 μL PfuTurbo (registered trademark) DNA polymerase (2.5 units/μL) (Stratagene) in a total volume of 50 μL. Cycling conditions were an initial denaturation step of 5 min at 94°C, followed by 35 cycles of 94°C for 20 sec, 50°C for 30 sec and 72°C for 1 min with a last treatment of 72°C for 10 min and then finally storage at 4°C.

- 74 -

PCRs were performed in a Perkin Elmer GeneAmp PCR System 9600.

# 32P-Labelling of DNA Probes

DNA fragments (50 to 100 ng) were radioactively labelled with 50  $\mu$ Ci of [ $\alpha$ -<sup>32</sup>P]-dCTP using a Gigaprime kit (Geneworks). Unincorporated [ $\alpha$ -<sup>32</sup>P]-dCTP was removed by chromatography on a Sephadex G-50 (Fine) column.

#### 10 Plasmid Isolation

5

15

25

Single colonies were analyzed for inserts by growing in LB broth (Sambrook et al., 1989, supra) with appropriate antibiotic selection (e.g. 100 µg/mL ampicillin or 10 to 50 µg/mL tetracycline for binary vector constructs). Plasmid DNA was purified using the alkali-lysis procedure (Sambrook et al., 1989, supra) or using The WizardPlus SV minipreps DNA purification system (Promega) or Qiagen Plasmid Mini Kit (Qiagen). Once the presence of an insert had been determined, larger amounts of plasmid DNA were prepared from 50 mL overnight cultures using a QIAfilter Plasmid Midi kit (Qiagen).

#### 20 DNA Sequence Analysis

DNA sequencing was performed using the PRISM (trademark) Ready Reaction Dye Primer Cycle Sequencing Kits from Applied Biosystems. The protocols supplied by the manufacturer were followed. The cycle sequencing reactions were performed using a Perkin Elmer PCR machine (GeneAmp PCR System 9600). Sequencing runs were performed by the Australian Genome Research Facility at The Walter and Eliza Hall Institute of Medical Research (Melbourne, Australia).

Homology searches against Genbank, SWISS-PROT and EMBL databases were performed using the FASTA and TFASTA programs (Pearson and Lipman, 1988) or BLAST programs (Altschul et al., J. Mol. Biol. 215(3): 403-410, 1990). Percentage sequence

- 75 -

similarities were obtained using LALIGN program (Huang and Miller, Adv. Appl. Math. 12: 373-381, 1991) using default settings.

Multiple sequence alignments were produced using ClustalW (Thompson et al., Nucleic Acids Research 22: 4673-4680, 1994).

#### Petunia transformations

#### (a) Plant Material

10

15

5

Leaf tissue from mature plants of P. hybrida cv Mitchell (or Ba20 or Br140w) was treated in 1.88% w/v sodium hypochlorite for 2 minutes and then rinsed three times in sterile water. The leaf tissue was then cut into 25-50 mm<sup>2</sup> squares and precultured on MS media (Murashige and Skoog, *Physiol. Plant 15: 73-97*, 1962) supplemented with 1.0 mg/L  $\alpha$ -benzylaminopurine (BAP) and 0.1 mg/L  $\alpha$ -naphthalene acetic acid (NAA) for 24 hours under white fluorescent lights.

#### (b) Co-cultivation of Agrobacterium and Petunia Tissue

A. tumefaciens strain AGL0 containing a binary vector were maintained at 4°C on LB agar plates with 50 μg/mL tetracycline. A single colony was grown overnight in liquid LB medium containing 40 μg/mL tetracycline. The following morning 1-2 mL of the overnight culture was added to a fresh batch of 25 mL liquid LB medium and the culture was grown at 37°C with shaking until an absorbance reading at 650nm (A<sub>650</sub>) of 0.4 to 0.8 was reached. A final concentration of 5 x 10<sup>8</sup> cells/mL was prepared by dilution in liquid MS medium containing 50 μM acetosyringone and 3% w/v sucrose B5 vitamins (Gamborg et al., Exp. Cell Res. 50: 151-158, 1968). The leaf discs were dipped for 2 minutes into the inoculum and then blotted dry and placed on co-cultivation media for 5 days. The co-cultivation medium consisted of SH medium (Schenk and Hildebrandt, Can. J. Bot. 50: 199-204, 1972) supplemented with 0.05 mg/L kinetin and 1.0 mg/L 2,4-D.

- 76 -

#### (c) Recovery of transgenic petunia plants

After co-cultivation, the leaf discs were transferred to selection medium (MS medium supplemented with 3% w/v sucrose, 3 mg/L BAP, 0.2 mg/L IAA, 1 µg/L chlorsulfuron, 300 mg/L timentin and 0.3% w/v Gelrite Gellan Gum (Schweizerhall). Regenerating explants were transferred to fresh selection medium after 2 weeks.

Adventitious shoots which survive the chlorsulfuron selection are isolated and transferred to BPM containing 1  $\mu$ g/L chlorsulfuron and 300 mg/L timentin for root induction. All cultures are maintained under a 16 hour photoperiod (60  $\mu$ mol. m<sup>-2</sup>, s<sup>-1</sup> cool white fluorescent light) at 23  $\pm$  2°C. When roots reach 2-3 cm in length the transgenic petunia plantlets are transferred to autoclaved Debco 51410/2 potting mix in 8 cm tubes. After 4 weeks, plants are be replanted into 15 cm pots, using the same potting mix, and maintained at 23°C under a 14 hour photoperiod (300  $\mu$ mol. m<sup>-2</sup>, s<sup>-1</sup> mercury halide light).

15

10

5

#### Arabidopsis transformations

Arabisopsis thaliana ecotype WS-2 seeds were obtained from The University of Melbourne, Parkville, Melbourne, Australia.

20

Plant growth conditions and transformation of A. thaliana were as essentially as described by Clough and Bent, (Plant J., 16: 735-743, 1998) except that seeds from the transformed plants were selected on 100  $\mu$ g/mL chlorsulfuron when binary vectors containing the SuRB selectable marker gene were used for the transformation process.

25

30

#### **EXAMPLE 2**

#### Isolation of new colored-protein sequences from Heron Island coral

Coral samples were collected from Heron Island Reef flat, Queensland, Australia. These samples were viewed as whole tissue under a fluorescent microscope, as described herein.

-77 -

#### Assessment of fluorescence properties

Table 2 shows taxonomic relationships of GFP isolated from the phylum Cnidaria and comparison with one amino acid sequence of the invention (Aams2-pep; SEQ ID NO:88). Fluorescent properties were analysed using an Olympus fluorescent microscope (BH2 - RFL) with filter combinations, as shown in Table 3. Tables 4 and 5 show fluorescent properties of colors for different species of organisms from Anthozoa and Hydrozoa.

#### Total RNA isolation

10

15

5

Plating corals were ground with a mortar and pestle and branching corals were scrubbed with a toothbrush directly into cold solution D, as described in Chomczynski and Sacchi, 1987, supra. Solution D-comprising tissue was homogenized using a glass homogenizer and transferred to 1.5 ml eppendorf microcentrifuge tubes. A 10% w/v 2 M sodium acetate (pH 4) solution was added prior to phenol chloroform extraction and extracted material was precipitated overnight in isopropanol at -20°C. Pellets were resuspended in solution D, and precipitated again in isopropanol. Resulting pellets were dissolved in 3 mM EDTA and 50 mM sodium acetate (pH 5) to be finally precipitated and stored at -20°C in ethanol.

#### 20 cDNA construction

RNA isolated from collected coral tissue was used to prepare cDNA. cDNA were constructed using a directional cDNA synthesis kit from Clontech Laboratories (Palo Alto, CA, USA) herein incorporated by reference.

25

# 5' Forward primers for PCR amplification

**SEQ ID NO:1** 

POC FOR

30 TCC GTT ATC GCT AAA CAG ATG ACC TAC AAA

**SEQ ID NO:2** 

**POC 220** 

- 78 -

GGC GAC CAC AGG TTT GCG TGT

**SEQ ID NO:3** 

MSVIAT(FOR)

5

10

15

20

25

30

:

ATG AGT GTG ATC GCT ACA CAA

SEQ ID NO:1 was previously designed as a 5' (or forward primer) for PCR amplification of nucleic acids encoding coral pigment proteins. SEQ ID NO:1 was shown to anneal to nucleic acids encoding a polypeptide comprising amino acids, SVIAK (SEQ ID NO:5): Refer to Dove et al. (2001; supra) and International Patent Publication No. WO 00/46233.

SEQ ID NO:2 was originally designed as a 3' (or reverse primer) for PCR amplification of nucleic acids encoding coral pigment polypeptides as disclosed in WO 00/46233. In addition to annealing to a 3' region of the nucleic acid as intended, SEQ ID NO:2 also anneals to a 5' UTR region of pocilloporin from *Acropora aspera* as disclosed herein.

SEQ ID NO:3 is newly designed and synthesized based on sequence information from PCR amplification products using SEQ ID NO:1 and SEQ ID NO:2. The amplified products comprise 5' UTR nucleotide sequence that includes sequence encoding a novel amino terminal end for a polypeptide similar to, but distinct from, the polypeptide disclosed in International Patent Publication No. WO 00/46233. This novel polypeptide has an amino terminal end comprising amino acids (M)SVIAT (SEQ ID NO:6; Figure 3). Accordingly, SEQ ID NO:3 anneals to nucleic acids encoding a peptide comprising (M)SVIAT (SEQ ID NO:6). Although peptide sequences SVIAK (SEQ ID NO:5) and (M)SVIAT (SEQ ID NO:6) differ by only one amino acid, the corresponding nucleic acids only share 67% identity (12 nucleic acids of 18). Notably, SEQ ID NO:1 cannot be used to amplify sequences starting with the N-terminal peptide (M)SVIAT (SEQ ID NO:6), and SEQ ID NO:3 cannot be used to amplify sequences beginning with the SVIAK (SEQ ID NO:5) peptide.

- 79 -

#### 3' Reverse primers for PCR amplification

**SEQ ID NO:4** 

**POC 231** 

5 TTT GTG CCT TGA TTT GAC TCT

SEQ ID NO:2 was also used as a 3' reverse primier and is described above. SEQ ID NO:4 was designed to anneal to a 3' end of previously isolated pocilloporin from *Acropora aspera* (Dove *et al.* [2001; *supra*] and International Patent Publication No. WO 00/46233).

PCR amplification

10

15

20

25

PCR amplification was performed using a combination of the abovementioned SEQ ID NOs as described in more detail hereinafter. Hybaid PCR express (Hybaid PCR Express, Integrated Sciences, Australia) was used according to instructions provided therein. Amplification products were separated by gel electrophoresis on a 1.5% w/v agarose gel and nucleic acid bands comprising desired nucleic acids were visualized using standard methods. Agarose gel comprising the desired nucleic acids were gel purified and the gel purified nucleic acids were inserted by ligation into pGemT-vector (Promega, Madison, WI, USA) producing a recombinant vector.

The inserted nucleic acids were sequenced using T7 and SP6 primers, which flank the inserted nucleic acid (sequencing service provided by AGRF; University of Queensland, Australia). Sequencing of the insert was performed at least twice in both forward and reverse directions until ambiguities were resolved. The following sequences were sequenced in a single direction: Ce61-7sv-rep (SEQ ID NO:37); Ce61-5sv-rep (SEQ ID NO:35); PM1Csv-rep (SEQ ID NO:57); PM1Asv-rep (SEQ ID NO:55); Mi68Dms (SEQ ID NO:119); Acams-3 (SEQ ID NO:101).

Table 6 shows amino acid sequences within 5 Angstroms of the fluorphore which encode possible spectral variants of the polypeptides of the invention comprising an amino acid sequence SGIAT (SEQ ID NO:7), SVIVT (SEQ ID NO:8), SVSAT (SEO ID NO:9) or

- 80 -

(M)SVIAT (SEQ ID NO:6) at the amino terminal end. These amino acid sequences were translated from nucleic acid sequences derived by PCR using 5' primers: SEQ ID NO:2 (5' UTR) and SEQ ID NO:3 [(M)SVIAT]; and 3' primers: SEQ ID NO:2 and SEQ ID NO:4.

Table 7 shows amino acid sequences within 5 Angstroms of the fluorphore which encode possible spectral variants of the polypeptides of the invention comprising an amino acid sequence SVIAK (SEQ ID NO:5) at the amino terminal end. These amino acid sequences were translated from nucleic acid sequences derived by PCR using 5' primer SEQ ID NO:1 and 3' primer SEQ ID NO:2, and 3' SEQ ID NO:3.

10

15

20

25

30

## Polypeptide modelling

A 3-dimensional model of the polypeptides was used to predict those amino acids within 5 Angstroms of the fluorophore "QYG". These amino acids have potential to influence spectral properties (Tsien, 1998, supra and Dove et al., 2001, supra) and are shown in Tables 6 and 7. The amino acids which are predicted to be located within 5 Angstroms of the fluorophore correspond to amino acid residues 37, 39, 56-65 (which comprises the fluorophore QYG), 86, 88, 90, 104, 106, 115, 139, 141, 143, 156, 158, 171, 192, 194, 208, 209 and 210. Amino acid residue numbers refer to numbering beginning with amino terminal amino acids S-V-I as residues 1, 2 and 3, respectively.

Information in relation to amino acid residues within 5 Angstroms of the fluorophore and details of atomic contacts for the polypeptide disclosed in Table 4 of International Patent Publication No. WO 00/46233, may be useful with the polypeptides of the present invention. In Tables 6 and 7, "Type" refers to a grouping or class of common amino acids within 5 Angstroms of the fluorophore, and "\*" indicates an internal stop codon. "Name" refers to consensus sequence name from multiple repeat sequences.

Figure 9 lists many of the pigment polypeptides of the invention and indicates the amino acid residues that are located within 5 Angstroms of a fluorophore region of the polypeptide. In addition, those amino acids residue positions where variation is found

- 81 -

throughout the different polypeptides are shown. Variable amino acids indicated throughout the polypeptide may be significant, as they may interfere with polypeptide folding.

#### 5 Amino acid and nucleotide sequence comparisons

10

Figures 1 and 3 show amino acid sequences for polypeptides comprising amino terminal SVIAK (SEQ ID NO:5; Figure 1) and comprising (M)SVIAT (SEQ ID NO:6), SGIAT (SEQ ID NO:7), SVIVT (SEQ ID NO:8) and SVSAT (SEQ ID NO:9) at or near the terminal amino end (Figure 3). Aams-2.pep (SEQ ID NO:88) and Aams-4.pep (SEQ ID NO:90) are shown comprising additional amino acids at the amino terminal end. Alignments of the corresponding nucleotide sequences of the amino acid sequences shown in Figures 1 and 3 are set forth in Figures 2 and 4, respectively.

Polypeptides comprising five shared amino acid sequences SVIAK (SEQ ID NO:5), (M)SVIAT (SEQ ID NO:6), SGIAT (SEQ ID NO:7), SVIVT (SEQ ID NO:8) and SVSAT (SEQ ID NO:9) may be grouped accordingly. Additional common amino acids immediately adjacent to the abovementioned amino acids are shown below:

SVIAT QMTY KVYM SGT (SEQ ID NO:10);

SVIAT QMTY KVYM PGT (SEQ ID NO:11);

SVIAT QVTY KVYM SGT (SEQ ID NO:12);

SGIAT QMTY KVYM SGT (SEQ ID NO:13);

SVIVT QMTY KVYM SGT (SEQ ID NO:14);

SVSAT QMTY KVYM SGT (SEQ ID NO:15);

SVIAK QMTY KVNM SGT (SEQ ID NO:16);

SVIAK QMTY KVYM SDT (SEQ ID NO:17); and

SVIAK QMTY X1X2YX3 SGT (SEQ ID NO:18),

wherein  $X_1$ ,  $X_2$  and  $X_3$  may be any amino acid provided that  $X_1$  is not  $X_2$  is not  $X_3$  is not  $X_3$ .

- 82 -

Figure 5 shows an alignment of amino acid sequences comprising SVIAK (SEQ ID NO:5) at the amino terminus and a stop or termination codon at corresponding amino acid residue 14. The termination codon results from the addition of two nucleic acid residues. The resulting polypeptide is much different when compared with polypeptides lacking this termination codon. An alignment of the corresponding nucleic acid sequences is shown in Figure 6. These nucleic acids are approximately 40 nucleotide bases longer than those lacking the termination codon (Figure 6). The differences can be more redily seen by referring to Figure 7, which shows an alignment of one nucleic acid sequence comprising the termination codon (SEQ ID NO:169) and a nucleic acid sequence lacking the termination codon (SEQ ID NO:19).

Previously-disclosed SVIAK (SEQ ID NO:5)-containing proteins Aapat-1 (SEQ ID NO:181) and Aapat-2 (SEQ ID NO:182) are also included on an amino acid sequence alignment with many of the SVIAK (SEQ ID NO:5)-containing polypeptides of the present invention, in Figure 8. Shaded amino acid residues indicate amino acids unique to SEQ ID NO:181 and/or SEQ ID NO:182.

#### **EXAMPLE 3**

Isolation of new colored-protein sequences from Melbourne coral

#### Extraction and visualization of colored proteins from coral

Samples of various coral and algae were purchased from Water World Aquarium (Melbourne, Australia) and Coburg Aquarium (Melbourne, Australia). These included Goniopora sp. ("flower pot coral") [brownish tentacles with an iridescent green centre underwater], green Acropora sp. coral ("staghorn coral"), brown/light blue Porites sp. coral ("finger"), Sinularia sp. and Tubastrea sp. corals as well as deep blue and bright orange Corallimorphs (Discosoma sp.).

10

15

20

25

- 83 -

Small samples of each coral were incubated in 1 M sodium phosphate buffer pH 7.5 at 4°C. A sample of "purple algae" that was growing on dead coral (normally sold as "living rock") was also collected in buffer. After 48 h the *Acropora* sp. extract appeared yellowbrown in color, the *Porites* sp. extract appeared orange in color and the purple algae extract was a clear pink color.

When the extracts were exposed to UV light the *Acropora* sp. extract contained orange and blue fractions, the *Porites* sp. extract contained pink fractions and the "purple algae" extract was a bright orange color.

10

5

Goniopora sp. coral tips were extracted in 1 M Na phosphate buffer pH 7.5. After an overnight incubation at 4°C the extract was orange-pink under natural light and appeared orange under UV light. Fluorescent green fractions were also observed in the solid phase under UV light.

15

20

A 10  $\mu$ L sample of the crude extracts described above was electrophoresed through precast SDS PAGE gels (12% w/v resolving, 4% w/v stacking gel) (Ready Gels, Biorad) in a running buffer made of 25 mM Tris-HCl, pH 8.3, 192 mM glycine, 0.1% w/v SDS at 100V for 75 min. The crude protein extracts were either denatured by boiling in 10% v/v glycerol, 3% w/v SDS, 3%  $\beta$ -mercaptoethanol, 0.025% w/v bromophenol blue or loaded in their native state in 5% v/v glycerol, 0.04% w/v bromophenol blue. Standards included pre-stained Low Range markers (Biorad) which contained standard protein samples of 116 kDa, 80 kDa, 51.8 kDa and 34.7 kDa.

Prior to staining with Coomassie blue (0.25% (w/v) Coomassie Brilliant Blue, 45% (v/v) methanol, 10% (v/v) acetic acid), PAGE gels were examined under a hand-held UV transilluminator (BLAK-RAY, longwave UV lamp, model B100 AP, UVP Inc). The non-denatured crude protein extract from Goniopora sp. contained orange bands (running higher than 116 kD marker protein) and blue-green bands (running between 51.8 kD and 80 kD protein markers). The non-denatured crude protein extract from Porites sp. contained two orange bands under UV light at approximately the same position as that

- 84 -

from Goniopora sp (i.e. running higher than 116 kD marker protein). The non-denatured crude protein extract from Acropora sp. contained a single orange band under UV light at approximately the same position as that from Goniopora sp. (i.e. running higher than 116 kD marker protein) as well as a green band (running between 80 kD and 116 kD marker proteins).

These fluorescent bands were not observed in any of the denatured crude protein extracts. No protein bands were visible under natural light before Coomasie blue staining.

#### 10 Isolation of RNA and synthesis of cDNA from coral

Total RNA was isolated from the anthozoans *Acropora* sp., *Discosoma* sp., *Sinularia* sp. and *Tubastrea* sp. using RNeasy Plant mini kit (Qiagen) or the method of Turpen and Griffith (*Biotechniques 4:* 11-15, 1986).

15

5

Complementary DNA was synthesized using 1 µg total RNA, 1 µL DNase RQ1 RNase free (Promega), 1 µL 10 x buffer (final concentration: 40 mM Tris-HCl pH 8, 10 mM NaCl, 6 mM MgCl<sub>2</sub>, 10 mM CaCl<sub>2</sub>). The reactions were incubated at 37°C for 10 min then 65°C for 10 min. One microlitre (1 µg) of primer dT(17)Ad2Ad1 (SEQ ID NO:183) was then added and the reaction was boiled for 5 min and then incubated on ice for 5 min. 4 µL 5 x RT buffer, 2 µL 0.1 M DTT, 1 µL 10 mM dNTPS and 1 µL RNasin (Promega) were then added and the reaction was incubated at 50°C for 2 min. 1 µL (200 U) Superscript II reverse trancriptase (Gibco BRL) was then added and the reaction was incubated at 50°C for 1.5 h. The cDNA was purified using QIAquick PCR purification kit (Qiagen).

25

.0

3

30

#### PCR of colored protein sequences

Oligonucleotide primers "vispro-F1" (SEQ ID NO:184) and "vispro-R1" (SEQ ID NO:185) were designed to hybridize to the 5' and 3' ends of T7SP6BASPOC3 and T7SP6BASPOC4 sequences, respectively (International Patent Application No. PCT/AU00/00056). The primer "vispro-F1" (SEQ ID NO:184) contained a *BamHI* site for

- 85 -

cloning into the bacterial expression vector pQE-30 (Qiagen) and an AscI site with a translation initiating codon for cloning into binary vectors. The primer "vispro-R1" (SEQ ID NO:185) contains a PstI site for cloning into the bacterial expression vector pQE-30 and a PacI site with translation termination codon for cloning into binary vectors.

5

10

20

25

30

SEQ ID NO: 184 vispro-F1 (5' to 3')

AscI

BamHI

CAG GGCGCGCC ATG GGA TCC GTT ATC GCT AAA CAG ATG ACC

**SEQ ID NO:185** 

vispro-R1 (5' to 3')

PacI PstI

15 GGG TTA ATT AAG CTG CAG GGC GAC CAC AGG TTT GCG TG

stop N L Q L A V V P K R

Polymerase chain reactions were set up using 20 pmole vispro-F1 (SEQ ID NO:184) and 20 pmole vispro-R1 (SEQ ID NO:185) primers and 5 μL cDNA synthesized from coral RNA as template, 2.5 units HotStarTaq (trademark) DNA polymerase (Qiagen), 200 μM dNTP mix and 1 X PCR buffer (Qiagen) in a 50 μL reaction.

PCR conditions included a denaturation step at 95°C for 15 min, followed by 35 cycles of 94°C for 30 sec, 50°C for 30 sec and 72°C for 1 min with a final treatment at 72°C for 10 min followed by storage at 4°C.

PCR products were electrophoresed through a 1% w/v agarose gel. Products of ~700 bp were excised from the gel and purified using QIAEX II Gel Extraction Kit (Qiagen). Purified DNA was digested with BamHI and PstI restriction enzymes and re-purified using a QIAquick PCR purification Kit (Qiagen). The purified DNA was ligated with BamHI/PstI ends of the bacterial expression vector pQE-30 (Qiagen). Ligated DNA was transformed into Eschericia coli BL21-RIL, M15 (containing pREP4 (Qiagen)) or XL1-blue competent cells and plated onto Luria Broth (LB) agar plates containing 100 μg/mL

- 86 -

ampicillin. After overnight incubation at 37°C a colony lift on nylon membrane (DuPont/NEN) was taken and placed colony side up onto LB agar containing 100 μg/mL ampicillin and 1 mM IPTG. The plates were incubated overnight at 37°C or alternatively at room temperature for 2 nights. Blue and purple colored colonies that were visible under natural light were obtained from products originating from *Acropora* sp., *Discosoma* sp., *Sinularia* sp. and *Tubastrea* sp.

Cultures of the purple and blue colonies were initiated and incubated overnight at 37°C. Plasmid DNA was isolated and analyzed by restriction endonuclease digestion. Plasmid DNA isolated from purple colonies included pCGP2915 (A10 clone from *Acropora* sp.), pCGP2916 (A11 clone from *Acropora* sp.), pCGP2917 (A12 clone from *Acropora* sp.), pCGP2918 (A8 clone from *Acropora* sp.), pCGP2920 (D10 clone from *Discosoma* sp.), pCGP2922 (T3 clone from *Tubastrea* sp.), pCGP2924 (S3 clone from *Sinularia* sp.).

Plasmid DNA isolated from blue colonies included pCGP2919 (D1 clone from *Discosoma* sp.), pCGP2921 (T1 clone from *Tubastrea* sp.), pCGP2923 (S1 clone from *Sinularia* sp.).

See Figure 10 for all schematics of above mentioned plasmids.

# 0 Sequence analysis of cDNA clones

5

10

2

5

2

Complete sequence analysis of the cDNA clones contained in the pQE-30 vectors was generated using pQEprom (Qiagen) (SEQ ID NO:186), pQErev (Qiagen) (SEQ ID NO:187), Coral-R1 (SEQ ID NO:188), vispro-F1 (SEQ ID NO:184) and vispro-R1 (SEQ ID NO:185) as sequencing primers.

SEQ ID NO:186 pQEprom

CCC GAA AAG TGC CAC CTG

- 87 -

#### SEQ ID NO:187 pQErev

GTT CTG AGG TCA TTA CTG G

5 SEQ ID NO:188 Coral-R1

TCA GGG TAC TTG GTG AAT GG

Complete nucleotide sequences were generated from the:-

10 A8 cDNA clone from Acropora sp. contained in pCGP2918 (SEQ ID NO:189);

D10 cDNA clone from Discosoma sp contained in pCGP2920 (SEQ ID NO:191);

S3 cDNA clone from Sinularia sp contained in pCGP2924 (SEQ ID NO:193);

T3 cDNA clone from Tubastrea sp. contained in pCGP2922 (SEQ ID NO:195);

D1 cDNA clone from Discosoma sp. contained in pCGP2919 (SEQ ID NO:197);

15 S1 cDNA clone from Sinularia sp. contained in pCGP2923 (SEQ ID NO:199); and

T1 cDNA clone from Tubastrea sp. contained in pCGP2921 (SEQ ID NO:201).

The A8 nucleotide sequence contained a putative open reading frame of 669 bases which encodes a putative polypeptide of 223 amino acids (SEQ ID NO:190).

20

The D10 nucleotide sequence contained a putative open reading frame of 669 bases which encodes a putative polypeptide of 223 amino acids (SEQ ID NO:192).

The S3 nucleotide sequence contained a putative open reading frame of 669 bases which encodes a putative polypeptide of 223 amino acids (SEQ ID NO:194).

2

:5

The T3 nucleotide sequence contained a putative open reading frame of 669 bases which encodes a putative polypeptide of 223 amino acids (SEQ ID NO:196).

0 The D1 nucleotide sequence contained a putative open reading frame of 669 bases which

encodes a putative polypeptide of 223 amino acids (SEQ ID NO:198).

- 88 -

The S1 nucleotide sequence contained a putative open reading frame of 669 bases which encodes a putative polypeptide of 223 amino acids (SEQ ID NO:200).

The T1 nucleotide sequence contained a putative open reading frame of 669 bases which encodes a putative polypeptide of 223 amino acids (SEQ ID NO:202).

Nucleotide and amino acid sequence similarities were determined using LALIGN (Huang and Miller, 1991, *supra*). The sequences isolated from the four species of coral share high nucleic acid and amino acid sequence similarities (Table 8 and Table 9).

10

15

5

#### **EXAMPLE 4**

#### Colored protein expression from Heron Island coral cDNAs

For expression in bacteria, nucleotide sequences encoding CFMs were retrieved from pGEM-T cloning vector using a forward oligonucleotide primer consisting of the NotI restriction binding site, a ribosomal binding site, a spacer and 15 bases encoding the N-terminus of the protein and a reverse oligonucleotide primer encoding H6-tag (POC220-H6; POC220 is SEQ ID NO:2). PCR product was gel purified and diluted (x10) prior to cloning into PCRII-TOPO and transformed into Top 10 cells (Invitogen). Cells were induced with 0.5 mM IPTG, and protein was purified on Ni-columns (Pro-Bond, Invitogen) eluting with 50 mM, 200 mM, 350 mM and 500 mM Imidazole in PBS pH 6.0, prior to overnight dialysis against 50 mM Potassium Phosphate pH 6.65.

# Expression of examples of Type 1 peptides

25

30

20

Results of expressing sequences of type 1 (as defined in Tables 6 and 7 and in Figure 9) in bacteria are set forth in Table 10. Only non-identical sequences are shown. Several additional sequences, which are identical to those shown in the Table, are indicated at the top of the Table (i.e.: Acasv-D = PavsvB, etc.). Sequence alignment is taken from International Patent Publication No. WO 00/46233 and Dove et al. (2001; supra). Horizontal bars above the amino acid sequence indicate  $\beta$ -strands from GFP structure. The

- 89 -

chromophore "QYG" is shown in white type on black background. Amino acid differences in the sequences are grey-shaded.

The majority of type 1 sequences are deep blue with  $\lambda_{max}$  ranging from 589 nm to 593 nm. Naturally-occurring amino acid substitution L161P, as seen in RTms5 (SEQ ID NO:166) compared with Acasv-D (SEQ ID NO:30) leads to clear bacteria that no longer absorb within 520-600 nm range. Reverse substitution of P161L re-establishes the ability to absorb in this range. The alignment shows amino acids that appear to affect colour of protein and those that do not.

10

15

30

5

Absorption scans for examples of expressed type 1 sequences are shown in Figure 11. Extinction coefficients at  $\lambda_{max}$ , as shown in this and in subsequent Figures 12 and 13, are based on the method of Whitaker and Granum (1980, *supra*) for protein detection. Extinction coefficient variability is partly due to the state of protein maturation; similar variability has been demonstrated for DsRed (Baird *et al.*, *Proc. Natl. Acad. Sci. USA 97*: 11984-11989, 2000).

#### Expression of examples of Types 2 and 14 peptides

Results of expressing sequences of type 2 in bacteria are shown in Table 11. Again, only non-identical sequences are shown. Additional sequences, identical to those shown in the Table, are indicated at the top of the Table (i.e.: PMms-B = PMms-E = PPd57-4ms, etc.). The majority of type 2 sequences are pinky-purple with λ<sub>max</sub> ranging from 579 nm to 580 nm. Naturally-occurring amino acid substitution P15S leads to clear bacteria that no longer absorb within 520-600 nm range. Alignment shows amino acids that do not affect the colour of protein, although it was noted that some of these proteins had a greater tendency to aggregate and precipitate than did others.

Analogous results, following expressing of type 14 sequences in bacteria, are shown in Table 12. Only non-identical sequences are shown. Table formatting is the same as in Tables 10 and 11. The majority of type 14 sequences are pinky-purple, with  $\lambda_{max}$  ranging

- 90 -

from 579 nm to 579.5 nm. Alignment shows amino acids that do not affect the colour of protein. It was noted, however, that MisvF and MisvA, with AA147 = F, was more soluble at higher concentrations than at others.

The spectral properties of Type 2 and Type 14 sequences are similar. This may be driven by AA61, which is Ser in both of these cases as opposed to Cys in type 1 and Thr in type 6 sequences. Figures 12A and B show absorption scans for examples of expressed type 2 and type 14 sequences. As described above for type 1 sequences, observed extinction coefficient variability is partly due to the state of protein maturation.

10

15

#### Expression of examples of Type 6 peptides

Examples of Type 6 sequences were similarly expressed in bacteria. Again, only non-identical sequences are shown. In this case, the majority of sequences are blue-purple, with  $\lambda_{\text{max}}$  ranging from 583.5 nm to 585.5 nm. Alignment shows that naturally occurring amino acid substitutions V8M and/or T182P lead to colourless bacteria, as does G238E, and that substitutions at AA101 and AA147 have slight effect on  $\lambda_{\text{max}}$ . Results are shown in Table 13 (see over). The format is the same as for Tables 10, 11 and 12.

Figure 13 shows absorption scans for examples of expressed type 6 sequences. As already stated above, extinction coefficient variability is partly due to the state of protein maturation and similar variability has been demonstrated for DsRed (Baird et al. 2000).

#### Expression of examples of peptides of other Types

25

Results of bacterial expression of sequence types other than the major types 1, 2, 6 and 14, are shown in Table 14 (see over). Many of the sequences that failed to express blue-purple or pink proteins were isolated from cDNA in which this was not the predominant GFP homolog present.

- 91 -

#### **EXAMPLE 5**

#### Estimation of amount of total soluble protein for colored proteins

Raw phosphate buffer extracts of two colour morphs of *Acropora aspera* (a dark blue pigmented morph and a cream morph) were used in the determination of the colored protein proportion of total soluble protein. Two separate estimations were made - by absorption spectroscopy and by gel filtration (n=5; 95% confidence intervals, in each case). Results are set forth in Figures 14A/B.

Figure 14A shows an absorption scan of the two Acropora aspera morphs. Estimation of blue-purple pocilloporin concentration (Dove et al., 1995, supra; Dove et al., 2001, supra) per surface area of coral tissue is based on an extinction coefficient range of 50,000 - 100,000 M<sup>-1</sup>cm<sup>-1</sup>. Figure 14B shows the results for chromatograms of gel filtrated protein elution, determined from 235 nm and 280 nm chromatograms, applying the equation (235 nm -280 nm)/2.51 (Whitaker and Granum, 1980, supra). The total area under the graph provides a measure of the total soluble protein. Blue-purple pocilloporin concentration is based on the difference between areas under the blue and cream graphs in the range of pocilloporin elution (24 - 26.5 min). Notably the independent methods for blue-purple pocilloporin concentration give similar results.

20

25

30

5

10

15

#### **EXAMPLE 6**

#### Colored protein expression from Melbourne coral cDNAs

Colonies of coral cDNA clones isolated from *Discosoma* sp. (D2 (pCGP2925 (blue in color)), *Sinularia* sp. (S1, pCGP2923) and *Tubastrea* sp. (T1, pCGP2921, T3, pCGP2922) were grown overnight with shaking at 37°C in 2mL LB media containing 100 µg/mL ampicillin. One mL of the overnight culture was then used to inoculate 25 mL LB media containing 100 µg/mL ampicillin. This culture was then incubated at 37°C with shaking until the OD<sub>600</sub> was around 0.5. IPTG was added to a final concentration of 1 mM and the cultures were grown overnight with shaking at 37°C. Cells (10 mL) of the incubated cultures were pelleted by centrifugation at 2000 rpm for 10 min. The bacterial pellets and

- 92 -

supernatant of the D2 (pCGP2925), S1 (pCGP2923) and T1 (pCGP2921) were blue those of T3 (pCGP2922) were purple under natural light. Bacterial pellets were stored at -20°C.

Proteins contained in the supernatant of the cultures were concentrated using Centricon 30 spin columns (Amicon) according to the manufacturer's instructions. The final volume of each of the concentrated protein extract was ~200 µL.

5

10

Aliquots (8 μL) of the concentrated proteins derived from the supernatants of the cultures were electrophoresced through precast SDS PAGE gels (12% w/v resolving, 4% w/v stacking gel) (Ready Gels, BIORAD) in a running buffer made of 25 mM Tris-HCl, Ph 8.3, 192 mM glycine, 0.1% w/v SDS at 100V for 75 min. Standards included Biorad Prestained Broad Range markers which contained standard protein samples of 206 kDa, 119 kDa, 91 kDa, 51.4 kDa, 34.7 kDa, 28.1 kDa, 20.4 kDa and 7.2k Da.

- Samples were either denatured by boiling in 10% v/v glycerol, 3% w/v SDS, 3% β-mercaptoethanol (βME), 0.025% w/v bromophenol blue or denatured by boiling in 10% v/v glycerol, 3% w/v SDS, 0.025% w/v bromophenol blue or loaded in their native state in 5% v/v glycerol, 0.04% w/v bromophenol blue.
- Prior to staining with Coomassie blue, protein bands were examined under a hand-held UV transilluminator. No fluorescent bands were visible under UV light in any of the samples. However, under natural light a blue band running at the same position as the 28 kDa protein standard was visible in the concentrated protein sample from the D2 supernatant. Blue smears that extended between the 28 kDa and 51 kDa protein standards were visible under natural light in the non-denatured concentrated protein samples from T1 and S1 supernatants. A purple smear which extended between the 28 kDa and 51 kDa protein standards was visible under natural light in the non-denatured concentrated protein samples from the S3 supernatant. There were no bands observed under natural light in samples that were denatured by boiling. Staining the gel with Coomassie blue showed that the proteins produced co-migrated with a 25 kDa protein marker (Biorad Precision Broadrange Prestained Marker).

- 93 -

Cultures of (E. coli XL1-blue) coral cDNA clones from Discosoma sp. (D1 in pCGP2919), Sinularia sp. (S1 in pCGP2923) and Tubastrea sp. (T1 in pCGP2921 and T3 in pCGP2922) that had grown at 37°C overnight with shaking were used to inoculate 100 mL · LB media containing 100 µg/mL ampicillin and further incubated with shaking at 37°C until the OD<sub>600</sub> was ~ 0.5. IPTG was added to a final concentration of 1 mM and the cultures were grown overnight with shaking at 37°C. Proteins expressed by Tubastrea sp. clones (T1 and T3) were purified under native conditions using Ni-NTA Superflow resin (Qiagen; QIAexpressionist 03/97) as recommended by the manufacturer. The elution buffer was exchanged with 20 mM Tris-HCl pH 8.0 using Sephadex G-25 columns (NAP10; Pharmacia) as per the manufacturer's instructions. Proteins expressed by the Discosoma sp. clone D1 and the Sinularia sp. clone S1 were purified under native conditions using the Ni-NTA method (Qiagen; QIAexpressionist 03/97) except that protein was precipitated from cleared bacterial lysate using 65% isopropanol and centrifuged at 10,000 rpm, 4°C, 10 min. The colored pellet was resuspended in 20mM Tris-HCl pH 8.0.

10

15

20

25

The proteins encoded by the *Acropora* sp. A8 clone in pCGP2918, the *Discosoma* sp. D10 clone in pCGP2920, the *Sinularia* sp S3 clone in pCGP2924 and the *Tubastrea* sp. T3 clone in pCGP2922 were a purple color (Royal Horticultural Society Color Chart (RHSCC) 88A) when concentrated. The proteins from *Tubastrea* sp. T3 clone and the *Sinularia* sp. S3 clone had absorbance peaks at approximately 580 nm.

The proteins encoded by the *Discosoma* sp. D1 clone in pCGP2919 and the *Tubastrea* sp. T1 clone in pCGP2921 were a blue color (RHSCC 102A) when concentrated and absorbance peaks at approximately 595 nm. The protein encoded by *Sinularia* sp. S1 clone in pCGP2923 was a blue-purple color (RHSCC 90A) when concentrated and had an absorbance peak at approximately 590 nm.

- 94 -

# Amino acid sequence alignment

A multiple alignment of the encoded amino acid sequence of T1 (SEQ ID NO:202), D1 (SEQ ID NO:198), S1 (SEQ ID NO:200), A8 (SEQ ID NO:190), T3 (SEQ ID NO:196), D10 (SEQ ID NO:192) and S3 (SEQ ID NO:194) was produced using the Clustal W (1.4) program in MacVector (6.5.3; Oxford Molecular Group Plc, 1999) (Figure 15). The multiple alignment of encoded amino acids showed that there are only 16 amino acid positions that differed between proteins exhibiting blue, blue-purple and purple color. From this alignment there appear to be eight amino acid positions that may influence the color of the protein (Table 15).

The protein encoded by S1 (SEQ ID NO:200) has a color that is intermediate of the blue and purple proteins. The amino acid sequence alignment (Figure 15) showed that the S1 amino acid sequence contained four amino acid identities characteristic of blue proteins towards the amino-terminal end and four amino acid identities characteristic to purple proteins towards the carboxy-terminal end (Table 15). The substitution of one or more amino acids listed in Table 15 may influence the visible color characteristics of the protein.

# Alignment of Melbourne and Heron Island coral protein sequences

20

5

10

15

The amino acid sequences of the above seven polypeptides (SEQ ID NOs 190, 192, 194, 196, 198, 200 and 202) were compared with other SVIAK (SEQ ID NO:5)-containing polypeptides, as set forth in Figure 1. The resulting alignment is shown in Figure 16.

25

30

#### **EXAMPLE 7**

# Expression of colored proteins in an eukaryotic organism Saccharomyces cerevisiae

In order to observe whether the colored protein sequences were able to produce color in a eukaryotic cell, the colored protein cDNA clones T1 (SEQ ID NO:201) and A8 (SEQ ID

- 95 -

NO:189) were introduced into a yeast expression vector (pYE22m) (Tanaka et al., 1988, supra) and transformed into Saccharomyces cerevisiae strain G1315.

# Construction of pCGP3269 and pCGP3270 (T1 or A8 in pYE22m)

5

The plasmids pCGP3269 (Figure 17) and pCGP3270 (Figure 18) were constructed by cloning the T1 or A8 cDNA clones, respectively, in a sense orientation behind the yeast glyceraldehyde 3-phosphate dehydrogenase promoter of pYE22m (Tanaka et al., 1988, supra).

10

15

A forward primer (Kpn.6His.F; SEQ ID NO:203) was designed to amplify the colored protein sequences that would result in 6 x Histidine tag fused in-frame with the colored protein at the N-terminus and a KpnI restriction endonuclease recognition site at the 5' end. A reverse primer (T1/A8.Sal.R; SEQ ID NO:204) included a SalI restriction endonuclease recognition site at the 3' end

SEQ ID NO:203 Kpn.6His.F

KpnI

SalI

20 GCAT GGT ACC ATG AGA GGA TCG CAT CAC CAT CAC CAT CAC

M R G S H H H H H H

**SEQ ID NO:204** T1/A8.Sal.R

25

30

CTGA GTC GAC TCA CTG CAG GGC GAC CAC AGG TTT

The coding regions of T1 (SEQ ID NO:201) and A8 (SEQ ID NO:189) were amplified by PCR using the primers Kpn.6His.F (SEQ ID NO:203) and T1/A8.Sal.R (SEQ ID NO:204) and the plasmid DNA pCGP2921 (T1) (Figure 10) and pCGP2918 (A8) (Figure 10) as template. The ~700bp PCR products were purified using a QIAquick PCR purification kit

- 96 -

(Qiagen) and then digested with the restriction endonucleases KpnI and SalI. The KpnI/SalI digested products were finally purified using a QIAquick PCR purification kit (Qiagen) and subsequently ligated with the KpnI/SalI ends of the pYE22m yeast expression vector (Tanaka et al., 1988 supra) using a DNA Ligation Kit (Amersham) according to the manufacturer's recommendations. Correct insertion of the T1 or A8 cDNA clones into the yeast expression vector was confirmed by visualisation of colour of transformants that were selected by their ability to restore G-1315 to tryptophan prototrophy. The T1 clone in the yeast expression vector pYE22m (designated as pCGP3269) produced blue coloured colonies (RHSCC 101C) when introduced into the yeast strain G1315. The A8 clone in the yeast expression vector pYE22m (designated as pCGP3270) produced purple coloured colonies (RHSCC 82B) when introduced into the yeast strain G1315.

#### **EXAMPLE 8**

15

10

5

# Estimation of colored protein amounts produced by bacterial and yeast cultures

#### Quantitation of colored protein expression in Saccaryomyces cerevisiae

Pure cultures of yeast cells harbouring pCGP3269 (Figure 17) or pCGP3270 (Figure 18) were grown at 29°C for 48 hours in 100 mL of YEPD liquid broth (1% yeast extract, 2% bacto-peptone, 2% w/v glucose, pH5.0). The cultures were centrifuged at 2000 rpm for 15 min. The resulting pellets were blue (pCGP3270) and purple (pCGP3269). The His-tagged colored proteins were extracted under native conditions by first resuspending the pellets in 4 mL lysis buffer (50 mM NaH<sub>2</sub>PO<sub>4</sub>, pH 8.0, 300 mM NaCl, 10 mM imidazole, 5 mg/mL Yeast Lytic enzyme (IBN)) and incubated at 30°C for 1 hour. The solutions were sonicated on ice 10 times for 10 sec with 15 sec cooling between treatments. The lysates were then centrifuged at 10 000 rpm for 10 min and the supernatants (crude extract) collected. The His-tagged colored proteins were purified by nickel-nitrilotriacetic acid metal-affinity chromatography (Qiagen) as recommended by the manufacturer.

- 97 -

The protein content of the crude extracts and purified His-tagged colored proteins were measured using a Bio-Rad Protein Assay using 1, 3 and 5 μL aliquots of extracts as per the manufacturer's instructions (Bio-Rad Microassay Procedure). The absorbances at 595 nm were compared with bovine serum albumin (BSA) standard curves (0-10 μg/mL) to obtain estimations of protein concentrations.

Samples of crude extracts and a dilution series of known amounts of purified His-tagged colored protein were electrophoresed through precast SDS PAGE gels (12% w/v resolving, 4% w/v stacking gel) (Ready Gels, Biorad) as described in Example 3. The gels were then stained with Coomasie blue (0.25% (w/v) Coomassie Brilliant Blue, 45% (v/v) methanol, 10% (v/v) acetic acid) and the amounts of colored protein in the crude extracts were estimated by comparing the intensities of the stained bands with those of the purified Histagged colored protein dilution series. This allowed the estimation of expression of colored protein in yeast as a percentage of total soluble protein (Table 16).

15

20

10

5

#### Quantitation of colored protein expression in Escherichia coli

One mL of an overnight Escherichia coli XL1blue culuture harbouring the plasmid pCGP2921 (T1) (Figure 10) (Example 3) was used to inoculate 100 mL LB broth (containing 50 µg/mL ampicillin) and incubated 37°C with shaking at 200 rpm until the OD600 was between 0.5 - 0.7. Protein production was induced with the addition of IPTG to 1 mM and incubation overnight at 29°C with shaking at 200 rpm. The cells were pelleted by centrifugation at 2000 rpm for 15 min. The resulting pellet was blue. The pellet was resuspended in 4 mL lysis buffer (50 mM NaH<sub>2</sub>PO<sub>4</sub>, pH 8.0, 300 mM NaCl, 10 mM imidazole) and sonicated on ice 6 times for 10 sec with 15 sec cooling between treatments. The solution was centrifuged at 10 000 rpm for 10 min and the (crude extract) supernatant collected. The His-tagged colored protein (T1) was extracted under native conditions by nickel-nitrilotriacetic acid metal-affinity chromatography (Qiagen) as recommended by the manufacturer.

25

- 98 -

The protein content of the crude extract and purified His-tagged colored protein was measured using a Bio-Rad Protein Assay using 1, 3 and 5  $\mu$ L of extracts as per the manufacturers instructions (Bio-Rad Microassay Procedure). The absorbances at 595 nm were compared with BSA standard curves (0-10  $\mu$ g/mL) to obtain estimations of protein concentrations.

Samples of crude extract and a dilution series of known amounts of purified His-tagged colored protein were electrophoresed through SDS PAGE gels as per the crude extract from yeast cultures (as described above). The amounts of colored protein in the crude extracts were estimated by comparing the intensites of the stained bands with those of the purified His-tagged colored protein dilution series. This allowed the estimation of expression of colored protein in *E. coli* as a percentage of total soluble protein (Table 16).

#### **EXAMPLE 9**

# Expression of colored proteins in plants under the control of a constitutive promoter

# Construction of pCGP2756 (35S: MCS: 35S expression cassette)

5

10

15

Plasmid pCGP2756 (Figure 19) was constructed by cloning the multicloning site (MCS) (containing the rare restriction endonuclease sites PacI and AscI) from pNEB193 (New England Biolabs) into the CaMV35S expression cassette of pRTppoptcAFP (Wnendt et al., Curr Genet 25: 510-523, 1994). The plasmid pRTppoptcAFP was digested with EcoRI and XbaI to release 300 bp AFP (antifungal protein) insert and the 3.3kb vector containing the CaMV 35S expression cassette. The plasmid pNEB193 was digested with EcoRI and XbaI to release the 40 bp fragment containing the multicloning site. The 40 bp EcoRI/XbaI fragment from pNEB193 and the 3.3 kb vector containing the CaMV35 expression cassette from pRTppoptcAFP were isolated and purified using the QIAEX II Gel Extraction kit (Qiagen) and ligated together. The ligation was carried out using the Amersham ligation kit.

- 99 -

analysis (Sall, KpnI, BamHI, Xbal, AscI, PacI, HindIII/BamHI) of DNA isolated from ampicillin-resistant transformants.

#### Construction of pCGP2757 (35S: MCS: 35S binary vector)

5

Plasmid pCGP2757 (Figure 20) was constructed by cloning the CaMV35S expression cassette of pCGP2756 (described above) into the binary vector pWTT2132 (DNAP). The plasmid pCGP2756 was digested with PstI to release the 0.7 kb CaMV35S expression cassette containing the multicloning site from pNEB193. The 0.7 kb fragment was isolated and purified using the QIAEX II Gel Extraction kit (Qiagen) and ligated with PstI ends of pWTT2132 binary vector. Correct insertion of the fragment in a tandem orientation to the CaMV35S: surB cassette in pWTT2132 was established by restriction enzyme analysis (KpnI, Pacl/AscI, EcoRI, XbaI, PstI) of DNA isolated from tetracycline-resistant transformants.

15

10

PCR products of CFMs or colored proteins derived using the primers vispro-F1 (SEQ ID NO:184) and vispro-R1 (SEQ ID NO:185) or using any primers containing AscI and PacI restriction endonuclease recognition sites, can be digested with AscI and PacI and ligated with AscI/PacI ends of pCGP2757.

20

25

30

#### Construction of pCGP2765 (35S: A8: 35S binary)

Plasmid pCGP2765 (Figure 21) was constructed by cloning the A8 PCR clone amplified from Acropora sp. into the CaMV35S expression cassette contained in the binary vector of pCGP2757 (described above). The A8 PCR product generated using the vispro-F1 (SEQ ID NO:184) and vispro-R1 (SEQ ID NO:185) primers and cDNA synthesized from Acropora sp. total RNA as template (see Example 1), was digested with AscI and PacI. The ~0.7 kb fragment was isolated and purified using the QIAEX II Gel Extraction kit (Qiagen) and ligated with AscI/PacI ends of pCGP2757 binary vector. Correct insertion of the fragment in a sense orientation behind the CaMV35S promoter was established by restriction enzyme analysis (EcoRI, PstI, BstXI) of DNA isolated from tetracycline-resistant transformants.

- 100 -

# Construction of pCGP2769 (35S: D1: 35S binary) (Figure 22)

5

10

15

Plasmid pCGP2769 (Figure 22) was constructed by cloning the D1 PCR clone amplified from *Discosoma sp.* into the CaMV35S expression cassette contained in the binary vector of pCGP2757 (described above). The PCR product generated using the primers vispro-F1 (SEQ ID NO:184) and vispro-R1 (SEQ ID NO:185) and the template pCGP2919 (containing the D1 cDNA clone) was digested with *AscI* and *PacI*. PCR was carried out in 50 μL reactions with 200 μM dNTPs, 20 pmol vispro-F1 (SEQ ID NO:184), 20 pmol visproR1 (SEQ ID NO:185), 1 x Pfu buffer (Stratagene), 2.5 units Pfu trubo DNA Polymerase (Stratagene) and ~2ng pCGP2919 plasmid DNA as template. The ~0.7kb fragment was isolated and purified using the QIAEX II Gel Extraction kit (Qiagen) and ligated with *AscI/PacI* ends of pCGP2757 binary vector. Correct insertion of the fragment in a sense orientation behind the CaMV35S promoter was established by restriction enzyme analysis (*EcoRI*, *PstI*, *BstXI*, *BamHI*) of DNA isolated from tetracycline-resistant transformants.

# Construction of pCGP2770 (35S: S1: 35S binary) (Figure 23)

Plasmid pCGP2770 (Figure 23) was constructed by cloning the S1 PCR clone amplified from Sinularia sp. into the CaMV35S expression cassette contained in the binary vector of pCGP2757 (described above). The PCR product generated using the primers vispro-F1 (SEQ ID NO:184) and vispro-R1 (SEQ ID NO:185) and the template pCGP2923 (containing the S1 cDNA clone) was digested with AscI and PacI. PCR was carried out in 50 μL reactions with 200 μM dNTPs, 20 pmol vispro-F1 (SEQ ID NO:184), 20 pmol vispro-R1 (SEQ ID NO:185), 1 x Pfu buffer (Stratagene), 2.5 units Pfu trubo DNA Polymerase (Stratagene) and ~2 ng pCGP2923 plasmid DNA as template. The ~0.7 kb fragment was isolated and purified using the QIAEX II Gel Extraction kit (Qiagen) and ligated with AscI/PacI ends of pCGP2757 binary vector. Correct insertion of the fragment in a sense orientation behind the CaMV35S promoter was established by restriction enzyme

- 101 -

analysis (EcoRI, PstI, BstXI, BamHI) of DNA isolated from tetracycline-resistant transformants.

# Construction of pCGP2772 (35S; T1: 35S binary) (Figure 24)

5

10

15

25

30

Plasmid pCGP2772 (Figure 24) was constructed by cloning the T1 PCR clone amplified from *Tubastrea sp.* into the CaMV35S expression cassette contained in the binary vector of pCGP2757 (described above). The PCR product generated using the primers vispro-F1 (SEQ ID NO:184) and vispro-R1 (SEQ ID NO:185) and the template pCGP2921 (containing the T1 cDNA clone) was digested with *Asc*I and *Pac*I. PCR was carried out in 50 μL reactions with 200 μM dNTPs, 20 pmol vispro-F1 (SEQ ID NO:184), 20 pmol vispro-R1 (SEQ ID NO:185), 1 x Pfu buffer (Stratagent), 2.5 units Pfu trubo DNA Polymerase (Stratagene) and ~2 ng pCGP2921 plasmid DNA as template. The ~0.7 kb fragment was isolated and purified using the QIAEX II Gel Extraction kit (Qiagen) and ligated with *AscI/PacI* ends of pCGP2757 binary vector. Correct insertion of the fragment in a sense orientation behind the CaMV35S promoter was established by restriction enzyme analysis (*EcoRI*, *PstI*, *BstXI*, *BamHI*) of DNA isolated from tetracycline-resistant transformants.

# 20 Construction of pCGP2926 (35S:His T1: 35S binary)

A histidine-tagged version of T1 was also produced for expression in the CaMV 35S gene expression cassette. The expression of this modified version of T1 will allow for a way of easily concentrating the expressed T1 protein to calculate the amount being produced in plants.

The RGS-His epitope was created by ligation of the 2 complementary primers TICS-His-FWD (SEQ ID NO:227) and TICS-His-REV (SEQ ID NO:228). This ligation resulted in a fragment containing the sequences to a prokaryotic ribosome binding site (RBS), a translational initiation consensus sequence (TICS) (for optimal translation in plants), the RGS-His epitope (consisting of sequences that encode the amino acids RGSHHHHHHH) and

- 102 -

overhanging AscI (at 5' end) and BamHI (at 3' end). This AscI/BamHI fragment was ligated with AscI/BamHI ends of plasmid pCGP2781 (Figure 32). Correct ligation of the insert into pCGP2781 was established by restriction enzyme analysis of DNA isolated

from tetracycline-resistant transformants. The plasmid was designated as pCGP2926 (Figure 44).

SEQ ID NO:227 TICS-His-FWD (5' to 3')

CGCGCC AAGGAGATAT AACA ATG AGA GGA TCG CAT CAC CAT CAC CAT CAC G

RBS TICS M R G S H H H H H H H H H RGS-His epitope

SEQ ID NO:228 TICS-His-REV (5' to 3')

GATCC GTG ATG GTG ATG GTG ATG CGA TCC TCT CAT TGTT ATATCTCCTT GG

RGS-His epitope TICS RBS

#### A. tumefaciens transformations

The plasmids pCGP2772 and pCGP2765 were introduced into the Agrobacterium tumefaciens strain AGL0 by adding 5 μg of plasmid DNA to 100 μL of competent AGL0 cells prepared by inoculating a 50 mL LB culture and growing for 16 hours with shaking at 28°C. The cells were then pelleted and resuspended in 0.5mL of 85% v/v 100 mM CaCl<sub>2</sub>/15% v/v) glycerol. The DNA-Agrobacterium mixture was frozen by incubation in liquid N<sub>2</sub> for 2 minutes and then allowed to thaw by incubation at 37°C for 5 minutes. The DNA/bacterial mix was then placed on ice for a further 10 minutes. The cells were then mixed with 1 mL of LB (Sambrook et al., 1989, supra) media and incubated with shaking for 16 hours at 28°C. Cells of A. tumefaciens carrying pCGP2772 and pCGP2765 were selected on LB agar plates containing 50 μg/mL tetracycline. The presence of pCGP2772 and pCGP27765 were confirmed by restriction enzyme analysis of DNA isolated from the tetracycline-resistant transformants.

- 103 -

#### **EXAMPLE 10**

## Spatial and temporal expression of colored proteins in plants

The use of constitutive promoters such as CaMV35S can be used to direct expression of CFM or colored proteins throughout the whole plant and may be useful in cases where a novel phenotype is sought with respect to the whole plant. However in some cases novel color is sought in specific tissues such as floral, seeds, leaves, fibre (e.g. cotton fibre), stems, roots, pollen, etc. In these cases tissue-specific promoters can be used to target expression of CFM or colored proteins to specific tissues. There are many cases in the literature, which describe the use of promoters to direct spatial and temporal expression. These promoters include, but are not limited to, the examples of a seed specific promoters (Song et al., Journal of Cotton Science 4: 217-223, 2000), leaf and chlorophyll containing tissue specific promoters (Song et al., 2000, supra), and tuber specific promoters (Rocha-Sosa et al., EMBO J 8: 23-29, 1989).

15

10

5

#### Isolation of Rose CHS promoter

A rose genomic DNA library was prepared from Rosa hybrida cv. Kardinal.

20 The rose library was screened with rose CHS cDNA clone.

A 6.6kb fragment upstream from the translational initiation site was cloned into pBluescript KS (-) (Stratagene) and the plasmid designated pCGP1114.

The plasmid pCGP1114 was digested with *Hind*III and *Eco*RV to release a ~2.7-3.0kb fragment which was purified using a Bresaclean kit (Geneworks) and ligated with *Hind*III/SmaI ends of pUC19 (New England Biolabs). Correct insertion of the Rose CHS promoter fragment was established by restriction enzyme analysis of DNA isolated from ampicillin-resistant transformants. The resulting plasmid was designated as pCGP1116 (Figure 25).

# - 104 -

# Construction of pCGP3255 (Rose CHS 5': 35S 3' pre-binary)

10

15

20

The plasmid pCGP3255 (Figure 26) was constructed by replacing the CaMV 35S promoter in the binary vector pCGP2757 with the Rose CHS promoter fragment from pCGP1116. Plasmid pCGP1116 was initially digested with *Hind*III. The overhanging 5' ends were filled-in using DNA polymerase (Klenow fragment) (Promega) according to the manufacturer's recommendation. The linearized vector was then digested with Asp718 to release a ~2.7kb rose CHS promoter fragment. The plasmid pCGP2757 was initially digested with *Sal*I. The overhanging 5' ends were filled-in using DNA polymerase (Klenow fragment) (Promega) according to the manufacturer's recommendation. The *Sal*I digested pCGP2757 was then digested with Asp718 to release the ~19kb binary vector fragment and the CaMV 35S promoter fragment. The *Sal*I (filled-in)/Asp718 ~19kb vector fragment was purified using QIAEX II Gel Extraction kit (Qiagen) and ligated with the *Hind*III (filled-in)/Asp718 ends of the rose CHS promoter fragment. Correct insertion of the rose CHS promoter was established by restriction enzyme analysis (*BgI*II, *Pst*I, *EcoRI*, *Hind*III, *Xba*I, *EcoRV*) of DNA isolated from tetracycline-resistant transformants.

PCR products of CFMs or colored proteins derived using the primers vispro-F1 (SEQ ID NO:184) and vispro-R1 (SEQ ID NO:185) or using any primers containing AscI and PacI restriction endonuclease recognition sites, can be digested with AscI and PacI and ligated with AscI/PacI ends of pCGP3255.

# Construction of pCGP2782 (Rose CHS: T1: 35S 3' binary)

- The plasmid pCGP2782 (Figure 27) was constructed by inserting the cDNA of the T1 coral protein contained in pCGP2921 (Example 1) behind the Rose CHS promoter contained in pCGP3255.
- The PCR product generated using the primers vispro-F1 (SEQ ID NO:184) and vispro-R1 (SEQ ID NO:185) and the template pCGP2921 (containing the T1 cDNA clone) was digested with AscI and PacI. PCR was carried out in 50 µL reactions with 200 µM dNTPs,

- 105 -

20 pmol vispro-F1 (SEQ ID NO:184), 20 pmol vispro-R1 (SEQ ID NO:185), 1 x Pfu buffer (Stratagene), 2.5 units Pfu trubo DNA Polymerase (Stratagene) and ~2ng pCGP2921 plasmid DNA as template. The resulting product was purified using QIAquick Gel Extraction (Qiagen) and ligated with AscI/PacI ends of pCGP3255. Correct insertion of the T1 coding region behind the Rose CHS promoter was established by restriction endonuclease digestion (HindIII, EcoRI, PstI, XbaI, BstX1) of tetracycline-resistant transformants.

#### Construction of pCGP2773 (Rose CHS: D1: 35S 3' binary)

10

5

The plasmid pCGP2773 (Figure 28) was constructed by inserting the cDNA of the D1 coral protein (Example 1) contained in pCGP2919 behind the Rose CHS promoter contained in pCGP3255.

The PCR product generated using the primers vispro-F1 (SEQ ID NO:184) and vispro-R1 (SEQ ID NO:185) and the template pCGP2919 (containing the D1 cDNA clone) was digested with AscI and PacI. The PCR product generated using the primers vispro-F1 (SEQ ID NO:184) and vispro-R1 (SEQ ID NO:185) and the template pCGP2919 (containing the D1 cDNA clone) was digested with AscI and PacI. PCR was carried out in 50 μL reactions with 200 μM dNTPs, 20 pmol vispro-F1 (SEQ ID NO:184), 20 pmol vispro-R1 (SEQ ID NO:185), 1 x Pfu buffer (Stratagene), 2.5 units Pfu trubo DNA Polymerase (Stratagene) and ~2ng pCGP2919 plasmid DNA as template. The resulting fragment was purified using QIAquick Gel Extraction (Qiagen) and ligated with AscI/PacI ends of pCGP3255. Correct insertion of the D1 coding region behind the Rose CHS promoter was established by restriction endonuclease digestion (HindIII, EcoRI, PstI, XbaI) of tetracycline-resistant transformants.

- 106 -

## Construction of pCGP2774 (Rose CHS: S1: 35S 3' binary)

The plasmid pCGP2774 (Figure 29) was constructed by inserting the cDNA of the S1 coral protein (Example 1) contained in pCGP2923 behind the Rose CHS promoter contained in pCGP3255.

The PCR product generated using the primers vispro-F1 (SEQ ID NO:184) and vispro-R1 (SEQ ID NO:185) and the template pCGP2923 (containing the S1 cDNA clone) was digested with AscI and PacI. The PCR product generated using the primers vispro-F1 (SEQ ID NO:184) and vispro-R1 (SEQ ID NO:185) and the template pCGP2923 (containing the S1 cDNA clone) was digested with AscI and PacI. PCR was carried out in 50 μL reactions with 200 μM dNTPs, 20 pmol vispro-F1 (SEQ ID NO:184), 20 pmol vispro-R1 (SEQ ID NO:185), 1 x Pfu buffer (Stratagene), 2.5 units Pfu trubo DNA Polymerase (Stratagene) and ~2ng pCGP2923 plasmid DNA as template. The resulting fragment was purified using QIAquick Gel Extraction (Qiagen) and ligated with AscI/PacI ends of pCGP3255. Correct insertion of the S1 coding region behind the Rose CHS promoter was established by restriction endonuclease digestion (HindIII, EcoRI, PstI, XbaI) of tetracycline-resistant transformants.

20

25

10

15

#### **EXAMPLE 11**

# Targeting of colored proteins to increase expression in plants

The levels of some CFMs or colored proteins produced in the cytosol of cells may have to be elevated in order to impart a visible color or a phenotype with commercial value. It is expected that targeting the CFM or colored proteins to different organelles within transgenic cells will significantly increase CFM or colored protein levels. The increased accumulation of transgene products by targeting to organelles has been demonstrated previously. For example, see Table 17.

30 It is also expected that plastid transformation of *Arabidopsis*, carnation, rose or other plant species will significantly increase CFM or colored protein levels. Increased accumulation

- 107 -

of transgene products by plastid transformation has been demonstrated previously. For example, see Table 18.

# Cloning of the chloroplast/plastid transit peptide sequence from tobacco

5

30

CFMs or colored proteins may be targeted to plastids with the inclusion of N-terminal plastid or chloroplast targeting peptides.

The 57 amino acid transit peptide of small subunit (SSU) of ribulose biphosphate carboxylase from *Nicotiana sylvestris* (Pinck et al., Biochimie 66: 539-545, 1984) was selected to target coral colored proteins to plastids of transgenic Arabidopsis, carnation, rose or other plant species.

The primers TSSU-Fnew (SEQ ID NO:205) and TSSU-R (SEQ ID NO:206) were used to amplify the tobacco chloroplast transit-peptide coding region using the plasmid pCGN5075 (Calgene) as template.

#### SEQ ID NO:205 TSSU-Fnew

20 CAG GGCGCCC AAGGAGATAT AACA ATG GCT TCC TCA GTT CTT TCC

AscI RBS TICS M A S S V L S

#### SEQ ID NO:206 TSSU-R

25 CACT GGATCC GCA TTG CAC TCT TCC GCC GTT GC

BamHI C Q V R G G N

TSSU-Fnew (SEQ ID NO:205) contains an AscI site for cloning into 35S and Rose CHS expression vectors, a prokaryotic ribosomal binding site (RBS) for bacterial expression and a plant translational initiation context sequence (TICS) for improved translation in plants. TSSU-R (SEQ ID NO:206) contains a BamHI site to allow the cloning of the transit

- 108 -

peptide in frame with coral colored protein sequences produced using vispro-F1 (SEQ ID NO:184) and vispro-R1 (SEQ ID NO:185) primers.

PCR conditions included 1 μL TSSU-Fnew (20 pmol/μL) (SEQ ID NO:205), 1 μL TSSU-R (20 pmol/μL) (SEQ ID NO:206), 5 μL 10 x pfu buffer (Stratagene), ~20ng pCGN5075 plasmid DNA as template, 1 μL 10mM dNTP mix, 0.5 μL Pfu turbo DNA polymerase (2.5 U/μL) (Stratagene) in a 50 μL reaction. The cycling conditions were 94°C for 5 minutes, followed by 35 cycles of 94°C for 30 min, 50°C for 30 min and 72°C for 60 min, and a final incubation at 72°C for 10 min. After completion of the PCR the products were stored at 4°C. PCR products were purified using a QIAquick PCR purification Kit (Qiagen) and cloned into pUC18 SmaI vector (Pharmacia/Amersham). The resulting plasmid was designated pCGP2783. The sequence of the transit peptide (TSSU) was confirmed by sequencing across both strands.

# 15 Construction of pCGP2780 (35S expression binary with unique BamHI site)

5

10

20

25

Plasmid pCGP2780 (Figure 30) was constructed by removing a ~290bp SaII fragment from pCGP2757. The plasmid pCGP2757 was digested with SaII to release a ~290bp fragment and ~19kb binary vector. The ~19kb binary vector was isolated and purified using the QIAEX II Gel Extraction kit (Qiagen) and self-ligated using the Amersham Ligation Kit. Correct religation of the SaII ends was established by restriction enzyme analysis (PvuII, BamHI, SaII) of DNA isolated from tetracycline-resistant transformants.

# Construction of pCGP2784 (35S expression pre-binary containing plastid transit peptide)

The plasmid pCGP2784 (Figure 31) was constructed by inserting the chloroplast transit peptide from tobacco contained in pCGP2783 into the binary vector pCGP2781.

Plasmid pCGP2783 was digested with AscI and BamHI to release the ~0.2 kb TSSU 30 fragment. The 0.2kb TSSU fragment was isolated and purified using the QIAEX II Gel Extraction kit (Qiagen) and ligated with AscI/BamHI ends of pCGP2781 binary vector.

- 109 -

Correct insertion of the transit peptide in frame and upstream of the T1 coding sequence was established by restriction enzyme analysis (*EcoRI*, *PstI*, *XbaI*, *AscI/PacI*) of DNA isolated from tetracycline-resistant transformants.

PCR products of CFMs or colored proteins derived using the primers vispro-F1 (SEQ ID NO:184) and vispro-R1 (SEQ ID NO:185) or using any primers containing BamHI and PacI restriction endonuclease recognition sites, can be digested with BamHI and PacI and ligated with BamHI/PacI ends of pCGP2784. The coding region of the CFMs or colored proteins will then be in-frame with the plastid targeting peptide to allow expression of the proteins in the plastids or chloroplasts.

# Construction of pCGP2781 (35S: T1: 35S binary with unique BamHI site)

Plasmid pCGP2781 (Figure 32) was constructed by removing a ~290bp Salī fragment from pCGP2772. The plasmid pCGP2772 was digested with Salī to release a ~290bp fragment and ~19kb binary vector. The ~19kb binary vector was isolated and purified using the QIAEX II Gel Extraction kit (Qiagen) and self-ligated using the Amersham Ligation Kit. Correct religation of the Salī ends was established by restriction enzyme analysis (Pvulī, BamHī, Salī, Xbal) of DNA isolated from tetracycline-resistant transformants.

20

15

# Construction of pCGP2785 (35S: TSSU: T1: 35S binary)

The plasmid pCGP2785 (Figure 33) was constructed by inserting the chloroplast transit peptide from tobacco contained in pCGP2783 into the binary vector pCGP2781.

25

30

Plasmid pCGP2783 was digested with AscI and BamHI to release the ~0.2 kb TSSU fragment. The 0.2kb TSSU fragment was isolated and purified using the QIAEX II Gel Extraction kit (Qiagen) and ligated with AscI/BamHI ends of pCGP2781 binary vector. Correct insertion of the transit peptide in frame and upstream of the T1 coding sequence was established by restriction enzyme analysis (EcoRI, PstI, XbaI, AscI/PacI) of DNA isolated from tetracycline-resistant transformants.

- 110 -

### Construction of pCGP2787 (Rose CHS: TSSU: T1: 35S binary)

The plasmid pCGP2787 (Figure 34) was constructed by inserting the chloroplast transit peptide from tobacco contained in pCGP2783 (Example 11) into the binary vector pCGP2782 (Figure 27).

Plasmid pCGP2783 was digested with AscI and BamHI to release the ~0.2 kb TSSU fragment. The 0.2kb TSSU fragment was isolated and purified using the QIAEX II Gel Extraction kit (Qiagen) and ligated with AscI/BamHI ends of pCGP2782 binary vector. Correct insertion of the transit peptide in frame and upstream of the T1 coding sequence was established by restriction enzyme analysis of DNA isolated from tetracycline-resistant transformants (Figure 34)

## 15 Targeting of CFMs or colored proteins to endoplasmic reticulum

CFMs or colored proteins are targeted to endoplasmic reticulum with the inclusion of N-terminal endoplasmic reticulum (ER) targeting peptides and C-terminal ER retaining signals.

20

25

30

5

10

The Arabidopsis thaliana basic chitinase N-terminal signal sequence was isolated to target CFMs and colored proteins to the ER (Haseloff et al., 1997, supra). To retain the proteins in the ER an HDEL peptide sequence was generated to be cloned in at the 3' end of the coding region (Haseloff et al., 1997, supra). These ER-targeting and ER-retention signals are used to increase levels of CFMs and colored protein in transgenic Arabidopsis, carnation, rose or other plant species.

The plasmid pBIN35Sm-GFP4-ER (Haseloff et al., 1997, supra) (http://www.plantsci.cam.ac.uk/Haseloff/GFP/mgfp4.html) was used as the source of Arabidopsis thaliana basic chitinase N-terminal signal sequence and HDEL ER-retention signal.

- 111 -

A PCR based approach was used to generate AscI and BamHI sites flanking the N-terminal ER transit peptide sequence. The primers AscI-ER.F (SEQ ID NO:207) and ER-BamHI.R (SEQ ID NO:208) were used to amplify the N-terminal ER sequence contained in pBIN35Sm-GFP4-ER.

Primer AscI-ER.F (SEQ ID NO:207) contains an AscI site for cloning into 35S and Rose CHS expression binaries (see Examples 9 and 10), a prokaryotic ribosome binding site (RBS) to allow for bacterial expression and a plant translational initiation context sequence (TICS).

10

15

25

GCAT GGCGCGCC AAGGAGATAT AACA ATG AAG ACT AAT CTT TTT C

Ascl RBS TICS M K T N L F

SEQ ID NO: 208 ER-BamHI.R (5' to 3')

BamHI ECORI

20 GCAT GGA TCC GAA TTC GGC CGA GGA TAA TGA TAG
S G F E A S S L S L

PCR conditions included using 1ng plasmid pBIN35Sm-GFP4-ER template, 100 ng each of primers AscI-ER.F (SEQ ID NO:207) and ER-BamHI.R (SEQ ID NO:208), 2.5 μL 10 x pfu turbo buffer (Stratagene), 1 μL pfu turbo (Stratagene) in a total volume of 25 μL. Cycling conditions were an initial denaturation step of 5 min at 94°C, followed by 35 cycles of 94°C for 30 sec, 50°C for 30 sec and 72°C for 1 min with a last treatment of 72°C for 5 min and then finally storage at 4°C.

An expected product of ~100bp was amplified and purified using the QIAEX II Gel Extraction kit (Qiagen) according to procedures recommended by the manufacturer. The 100bp fragment was then cloned into pCR2.1 (Invitrogen) and the plasmid was designated

- 112 -

pCGP3256. The sequence of the N-terminal ER transit peptide fragment was confirmed by sequence analysis using the M13 reverse and M13 -20 primers.

# Construction of pCGP3257 (35S:ER:MCS:35S pre-binary)

5

10

15

20

The N-terminal ER transit peptide fragment was cloned downstream of the 35S promoter contained in the pre-binary pCGP2780 (Figure 30) to produce pCGP3257 (Figure 35). Plasmid pCGP3256 was digested with AscI and BamHI to release the ~100bp N-terminal ER transit peptide fragment. The fragment was isolated and purified using QIAEX II Gel Extraction kit (Qiagen) and ligated with AscI/BamHI ends of pCGP2780. Correct insertion of the N-terminal ER transit peptide fragment was established by restriction endonuclease analysis of DNA isolated from tetracycline-resistant transformants.

PCR products of CFMs or colored proteins derived using the primers vispro F1 (SEQ ID NO:185) and CP-HDEL-PacI.R (described in this Example below) can be digested with BamHI and PacI and ligated with BamHI/PacI ends of pCGP3257. The coding region of the CFMs or colored proteins will be under the control of the CaMV 35S promoter and inframe with the ER transit targeting peptide to allow targeting of the proteins to the ER. The coding region of the CFMs or colored proteins will also contain the HDEL sequence at the C-terminal end to allow retention of the proteins in the ER.

# Construction of pCGP3259 (35S: ER: T1.HDEL: 35S binary)

25

The coding sequence of the colored protein T1 was amplified by PCR using the primers vispro-F1 (SEQ ID NO:184) and CP-HDEL-PacI.R (SEQ ID NO:209) and the plasmid pCGP2779 as template. The primer CP-HDEL-PacI.R was designed to include a PacI site with a translational termination codon for cloning into the binary vectors described in this specification, a HDEL peptide sequence in-frame with the colored protein sequence and a PstI site for cloning into the bacterial expression vector pQE-30 (Qiagen).

- 113 -

SEQ ID NO:209 CP-HDEL-PacI. R (5' to 3')

5

10

PacI

GATCTTAAT TAA AGC TCA TCA TGC TGC AGG GCG ACC ACA GGT TTG C

\* L B D H Q L A V V P K

PCR conditions included using 2ng plasmid pCGP2779 as template, 100ng each of primers vispro-F1 (SEQ ID NO:184) and CP-HDEL-PacI.R (SEQ ID NO:209), 2 μL 10 mM dNTP mix, 5 μL 10 x PfuTurbo (registered trademark) DNA polymerase buffer (Stratagene), 0.5 μL PfuTurbo (registered trademark) DNA polymerase (2.5 units/μL) (Stratagene) in a total volume of 50 μL. Cycling conditions were an initial denaturation step of 5 min at 94°C, followed by 35 cycles of 94°C for 20 sec, 50°C for 30 sec and 72°C for 1 min with a last treatment of 72°C for 10 min and then finally storage at 4°C.

The resulting ~700bp product was digested with BamHI and PacI, isolated and purified using QIAEXII Gel Extraction kit (Qiagen) and ligated with BamHI/PacI ends of pCGP3257. Correct insertion of the T1 coding region and HDEL sequence in-frame with the ER transit peptide sequence under the control of the 35S promoter was established by restriction endonuclease analysis (BamHI, EcoRI, AscI, PacI) of DNA isolated from tetracycline-resistant transformants. The resulting plasmid was designated pCGP3259 (Figure 36).

# Construction of pCGP3262 (RoseCHS:ER:MCS:35S pre-binary)

The N-terminal ER transit peptide fragment was cloned downstream of the Rose CHS promoter contained in the pre-binary pCGP3255 to produce pCGP3262 (Figure 37). Plasmid pCGP3256 was digested with AscI and BamHI to release the ~100bp N-terminal ER transit peptide fragment. The fragment was isolated and purified using QIAEX II Gel Extraction kit (Qiagen) and ligated with AscI/BamHI ends of pCGP3255. Correct insertion of the N-terminal ER transit peptide fragment was established by restriction endonuclease analysis of DNA isolated from tetracycline-resistant transformants.

- 114 -

PCR products of CFMs or colored proteins derived using the primers vispro-F1 (SEQ ID NO:184) and CP-HDEL-PacI.R (SEQ ID NO:209) can be digested with BamHI and PacI and ligated with BamHI/PacI ends of pCGP3262. The coding region of the CFMs or colored proteins will be under the control of the Rose CHS promoter and in-frame with the ER transit targeting peptide to allow targeting of the proteins to the ER. The coding region of the CFMs or colored proteins will also contain the HDEL sequence at the C-terminal to allow retention of the proteins in the ER of floral tissues.

# Construction of pCGP3263 (Rose CHS:ER: T1-HDEL:35S binary)

10

20

5

The coding sequence of the colored protein T1 was amplified by PCR using the primers vispro-F1 (SEQ ID NO:184) and CP-HDEL-PacI.R (SEQ ID NO:209) and the plasmid pCGP2779 as template.

15 PCR conditions were as described above for construction of pCGP3259.

The resulting ~700bp product was digested with BamHI and PacI, isolated and purified using QIAEX II Gel Extraction kit (Qiagen) and ligated with BamHI/PacI ends of pCGP3262. Correct insertion of the T1 coding region and HDEL sequence in-frame with the ER transit peptide sequence under the control of the Rose CHS promoter was established by restriction endonuclease analysis (BamHI, EcoRI, AscI, PacI) of DNA isolated from tetracycline-resistant transformants. The resulting plasmid was designated pCGP3263 (Figure 38).

A site predicting N-glycosylation was identified within the coloured protein T1 ('NDS' surrounding amino acid 107) (SEQ ID NO:202). This site is conserved among the colored
protein clones D1, D10, T1, T3, S3 and A8 and these include both purple and blue
varieties. Comparison of this region in sequences of other coloured and fluorescent
varieties in the GenBank database (e.g., asCP562, asFP499, Clavularia FP484, Discosoma
FP483 etc) indicate the presence of two alternative sequences in this position - QDS or
NDI. The first converts an asparagine residue (N) to a glutamine (Q) (a conservative

- 115 -

change given both residues are polar) and the second changes the serine (S) to an isoluecine (I) (a non conservative change from a polar to a non polar residue). Both naturally occurring sequence alternatives for this region of the protein were be performed separately. That is, mutation of the T1 sequence from NDS to QDS and a separate mutation from NDS to NDI.

The plasmid pCGP2921 (Figure 10) was used as a source of the coding sequence for T1 blue protein. A BamHI/HindIII fragment was isolated from pCGP2921 and cloned with BamHI/HindIII ends of pBluescript to produce pCGP3268. The GeneEditor in vitro Site Directed Mutagenesis Kit (Promega) was used following the manufacturer's instructions along with the following oligonucleotides (T1.N-Q N(AAT) > Q(CAG) SEQ ID NO:230) and T1.S-I S(TCC) > I(ATC) SEQ ID NO:231) to introduce the mutations in pCGP3268.

SEQ ID NO:230 T1.N-Q  $N(\underline{A}\underline{A}\underline{T}) > Q(\underline{C}\underline{A}\underline{G})$ 

15

10

5

GTG TGT ACT GTC AGC CAG GAT TCC AGC ATC CAA G
V C T V S O D S S I O

SEQ ID NO:231  $T1.S-IS(\underline{TC}C) > I(\underline{AT}C)$ 

20 CT GTC AGC AAT GAT ATC AGC ATC CAA GGC AAC

The resultant plasmids pCGP3271 and pCGP3272 containing the N107Q and S109I mutated forms of T1 blue protein in pBluescript were sequenced thoroughly to confirm the presence of the mutated sequence.

25

30

## Construction of pCGP3273 (pOE30:T1(N1070) and pCGP3274 (pOE30:T1(S1091)

E. coli expression of the mutated forms of T1 in pCGP3271 and pCGP3272 was necessary to determine if the mutations had any effect on the colour of the expressed protein. Thus, BamHI/HindIII fragments pCGP3271 and pCGP3272 were subcloned with BamHI/HindIII ends of pQE30. The resultant plasmids were designated pCGP3273 (T1- N107Q) and pCGP3274 (T1-S109I) and were expressed in E. coli as previously described (Example 3)

- 116 -

and 6) to determine the colour of the expressed protein. The protein expressed by the sequence encoded in pCGP3273 was found to retain the original colour of T1 as expressed by pCGP2921, while the protein expressed by pCGP3274 was not coloured. This suggested that the S109I mutation may have had a deleterious effect on the color of the protein. Investigation of this protein will provide information on the amino acids that are critical to color formation of colored proteins.

# Construction of pCGP3275 (35S: ER:T1(N1070).HDEL:35S binary) and pCGP3276 (35S: ER:T1(S109I).HDEL:35S binary)

10

The coding sequence of the coloured protein T1(N107Q) was amplified by PCR using the primers vispro-F1 (SEQ ID NO:184) and CP-HDEL-PacI.R (SEQ ID NO:207) and the plasmids pCGP3271 (described above) and pCGP3272 (described above) as template essentially as described in the construction of pCGP3259 (Example 11).

15

20

25

30

The resulting ~700bp products were digested with BamHI and PacI, isolated and purified using QIAEXII Gel Extraction kit (Qiagen) and ligated with BamHI/PacI ends of pCGP3257 (Figure 35). Correct insertion of the coding regions of T1(N107Q) and T1(S109I) and HDEL sequence in-frame with the ER transit peptide sequence under the control of the CaMV 35S promoter was established by restriction endonuclease analysis (BamHI, EcoRI, AscI, PacI, EcoRV) of DNA isolated from tetracycline resistant transformants. The resulting plasmids were designated pCGP3275 and pCGP3276.

# Construction of pCGP3277 (RoseCHS: ER:T1(N1070).HDEL:35S binary) and pCGP3276 (Rose CHS: ER:T1(S1091).HDEL:35S binary)

The coding sequence of the coloured protein T1(N107Q) was amplified by PCR using the primers vispro F1 (SEQ ID NO:184) and CP-HDEL-PacI.R (SEQ ID NO:207) and the plasmids pCGP3271 and pCGP3272 as template essentially as described in the construction of pCGP3259 (Example 11).

- 117 -

The resulting ~700bp products were digested with BamHI and PacI, isolated and purified using QIAEXII Gel Extraction kit (Qiagen) and ligated with BamHI/PacI ends of pCGP3262 (Figure 37). Correct insertion of the coding regions of T1(N107Q) and T1(S109I) and HDEL sequence in-frame with the ER transit peptide sequence under the control of the Rose CHS promoter was established by restriction endonuclease analysis (BamHI, EcoRI, AscI, PacI, EcoRV) of DNA isolated from tetracycline resistant transformants. The resulting plasmids were designated pCGP3277 and pCGP3278.

#### **EXAMPLE 12**

10

5

# Fusion proteins with GFP

# Construction of pCGP3258 (35S:T1/mGFP4:35S binary)

As a way of tracking the expression and localisation of the T1 coloured protein the T1 coding region was fused with the N-terminus of mGFP4 (Haseloff *et al.*, *PNAS 94*: 2122-2127, 1997).

The mGFP4 coding sequence was amplified using the primers PstI-mGFP4F (SEQ ID NO:210) and mGFP4-PacIR (SEQ ID NO:211) and pBIN35SmGFP4ER (Haseloff et al., 1997) as template. A ~700bp product was gel purified and then digested with the restriction endonucleases PstI and PacI. The T1 coding sequence was amplified using the primers visproF1-new (SEQ ID NO:212) and visproR1 (SEQ ID NO:185) and pCGP2779 as template.

25 SEQ ID NO:210 Pst-mGFP4F (5' to 3')

PstI linker sequences

GCAT CTG CAG GTC GCC ACC AGT AAA GGA GAA GAA CTT TTC AC

L Q V A T S K G E E L F

20

- 118 -

SEQ ID NO:211 mGFP4-PacIR

PacI

15

20

30

CTGA TTAATTAA TTA TTT GTA TAG TTC ATC CAT GCC ATG

5 \* K Y L B D M G H

SEQ ID NO:212 visproF1-new

AscI RBS TICS BamHI

10 CAG GGCGCGC AAGGAGATAT AACA ATG GGA TCC GTT ATC GCT AAA CAG ATG ACC

A ~700bp product was gel purified and then digested with the restriction endonucleases AscI and PstI.

The PstVPacI mGFP4 fragment was ligated with the AscVPstI T1 fragment. The resulting ligated fragment was then ligated with the AscVPacI ends of the binary vector pCGP3257 (Figure 35) to produce pCGP3258 (Figure 39). Correct insertion of the T1:mGFP4 fusion was established by restriction endonuclease analysis (BstX1, EcoRI, NcoI, PstI) of DNA isolated from tetracycline-resistant transformants. The resulting plasmid was designated pCGP3258 (Figure 39).

# Construction of pCGP3261 (35S:ER:T1:GFP: 35S binary)

An ER targeted version of the T1:mGFP4 fusion in pCGP3258 under the control of the CaMV 35S promoter was also prepared. This plasmid was designated pCGP3261 (Figure 45).

The T1:mGFP4 fusion was amplified using the primers vispro-F1 (SEQ ID NO:184) and mGFP4-HDEL-PacR (SEQ ID NO:229) and pCGP3258 (Figure 39) as template. A ~1.4kb product was gel purified and then digested with the restriction endonucleases BamHI and PacI. The resulting fragment was then ligated with BamHI/PacI ends of the binary vector

- 119 -

pCGP3257 (Figure 35) to produce pCGP3261 (Figure 45). Correct insertion of the T1:mGFP4 fusion was established by restriction endonuclease analysis (*BstX1*, *EcoRI*, *NcoI*, *PstI*, *AscI/PacI*, *XbaI*) of DNA isolated from tetracycline-resistant transformants. The resulting plasmid was designated pCGP3261 (Figure 45).

5

SEQ ID NO:229 mGFP4-HDEL-PacR (5' TO 3')

CTG ATT AAT TAA AGC TCA TCA TGT TTG TAT AGT TCA TCC ATG CCA TG

# 10 Construction of pCGP3260 (35S:ER:GFP: 35S binary)

An ER targeted version of the mGFP4 in pBIN35SmGFP4ER (Haseloff et al., 1997 supra) under the control of the CaMV 35S promoter and CaMV 35S terminator was prepared to use as a control for the binaries pCGP3258 (Figure 39) and pCGP3261 (Figure 45).

The plasmid pBIN35SmGFP4ER (Haseloff et al., 1997 supra) was initially digested with the restriction endonuclease SacI. The resulting overhang was repaired and the linearized vector was then digested with BamHI to release a ~0.7kb fragment containing the mGFP4 coding sequence. The resulting SacI(blunt)/BamHI mGFP4 fragment was gel purified and then ligated with BamHI/PacI (blunt) ends of the binary vector pCGP2780 (Figure 30).

Correct insertion of the mGFP4 coding sequence was established by restriction endonuclease analysis (EcoRI, NcoI, PstI, BamHI, XbaI) of DNA isolated from tetracycline-resistant transformants. The resulting plasmid was designated pCGP3260 (Figure 46).

25

30

#### **EXAMPLE 13**

## Reconstruction of color

In order to determine whether rose petals or plant material in general conatin proteases that may degrade colored proteins reconstructions of rose petal extracts with the T1 colored protein were set up.

- 120 -

Petals of Rosa hybrida cultivar Medeo are generally white to pale apricot. Expression of colored proteins in a white flower should allow visualisation of color when colored proteins are expressed in flowers.

- 5 One gram amounts of Medeo rose petals were ground in 500 µL water using a mortar and pestle. The resultant slurries were centrifuged at 14 000 rpm for 5 min in 1.5mL centrifuge tubes. The supernatants were collected and 100 µL of the extracts were aliquoted into the wells of a microtitre tray. Ten microlitres aliquots containing ~30 µg of His-tag purified T1 protein (purified as described in Example 8) were added to the Medeo extracts. In order to 10 determine whether the color of the colored protein is affected by pH, the pH of some of the reconstructions was modified by addition of NaOH so that the final pH was 7.0, 8.5 or 10.0. The pH of Medeo petal extract alone was pH 4.5 and 4.6. The pH of Medeo petal extract mixed with T1 protein was pH 5.2, 5.8 and 6.1. The color of reconstructions of Medeo petal extract mixed with T1 protein at pH 5.2, 5.8 and 6.1 was light blue (RHSCC 15 101 C/ RHSCC 115B). However the color at pH 7.0 and 8.5 was a pale blue-green (RHSCC 122C) and that at pH 10.0 was yellow. The colors were still evident after 5 hours incubation at room temperature as well as 48 hours at room temperature indicating that the colored protein was stable in petal extract.
- An interesting and unexpected observation was that the color of the T1 protein changed to yellow when in a high pH solution. Analysis of the conformation of the protein at this high pH provides information that allows for the design of targeted mutations to T1 or other colored protein sequence and thus allows for the production of a yellow color in a low to neutral pH environment such is found in plant cells. Alternatively random shuffling (US Patent No. 6, 132 970) using selections of the vast number of colored protein sequences isolated and then expressing these mutated versions in E. coli or yeast as described in Examples 3, 4, 6 and 7 will provide a means of selecting for altered or improved colors and/or brightness of the proteins expressed.

- 121 -

The flowers of *Petunia hybrida* cultivar Mitchell are white. Mitchell petal sections were incubated with the T1 protein to determine the color that would be produced in white petals upon production of the colored proteins. Petal sections (including part of the tube and limb) were incubated in 200 µL His-tag purified T1 protein (from *E. coli* cultures as described in Example 8) (6 mg/mL in 20 mM Tris HCl pH 8.0) and His-tag purified A8 protein (from yeast cultures as described in Example 8) (1 mg/mL in 20 mM Tris HCl pH 8.0). In both cases the colored proteins were taking up by the petal fragments within a few minutes as visualised by coloration of the cut surface of the petal. Incubation of white petals in the T1 protein solution resulted in petals of a pale blue (RHSCC 112D) color whereas incubation of white petals in the A8 protein solution resulted in a pale purple color in the petal tissue. This experiment showed that the protein is stable in petal tissue and that the color produced will not be masked or quenched by other plant compounds.

#### **EXAMPLE 14**

Expression of colored proteins in Arabidopsis

## Transformation of Arabidopsis

Construction of pCGP960 (35S:gus:ocs binary)

20

5

10

15

The binary vector pCGP960 was prepared to use as a control in plant transformation experiments. A CaMV35S:GUS:ocs3' expression cassette was isolated from pKIWI101 (Klee et al., Bio/Technology 3: 637-642, 1985) and inserted into the pWTT2132 (DNAP) binary vector backbone which contains a CaMV 35S:SuRB selectable marker gene.

25

30

The binary vectors pCGP2772 (Figure 24), pCGP2765 (Figure 21), pCGP3259 (Figure 36), pCGP2785 (Figure 33), pCGP3258 (Figure 39), pCGP2926 (Figure 44), pCGP3263 (Figure 38), pCGP2787 (Figure 34), pCGP2782 (Figure 27), pCGP960 (see above), pCGP3261 (Figure 45), pCGP3260 (Figure 46), pBINmGFP4ER (Haseloff et al., 1997, supra) were introduced into Agrobacterium tumefaciens strain AGL0 as described in Example 1.

- 122 -

Arabidopsis thaliana ecotype WS-2 was transformed with the above constructs using the floral dip method as mentioned in Example 1. Seeds from dipped plants were plated on selection and transgenic plants were allowed to grow until flowering. Plants can be allowed to self-fertilize to produce seed. The T2 seed can then be germinated on selection (e.g. 100 µg/mL chlorsulfuron selection for those transformed with a CaMV 35S: SuRB selectable marker gene) and allowed to grow to flowering. A number of the T2 generation would be expected to be homozygous for the introduced transgenes with the expectation that these plants would have increased coloured protein gene expression and protein production than the heterozygous parental lines.

#### Northern analysis

10

Leaves from a random selection of 2 events per construct (pCGP2772, pCGP2765, pCGP3259, pCGP2785, pCGP3258, pCGP3261, pCGP960, pBIN35Smgfp4ER, pCGP3260) were analysed for the presence of transcripts of the introduced T1 or A8 colored protein genes. Total RNA was isolated from these events using a Plant RNAeasy kit (QIAGEN) following procedures recommended by the manufacturer.

- 20 RNA samples (5 μg) were electrophoresed through 2.2 M formaldehyde/1.2% w/v agarose gels using running buffer containing 40 mM morpholinopropanesulphonic acid (pH 7.0), 5 mM sodium acetate, 0.1 mM EDTA (pH 8.0). The RNA was transferred to Hybond-N filters (Amersham) as described by the manufacturer.
- The RNA blot was initially probed with <sup>32</sup>P-labelled fragments of a BamHI/HindIII fragment isolated from pCGP2921 (T1) (Figure 10) (10<sup>8</sup> cpm/μg, 2 x 10<sup>6</sup> cpm/mL). Prehybridization (1 hour at 42°C) and hybridization (16 hours at 42°C) of the membrane were carried out in 50% v/v formamide, 1 M NaCl, 1% w/v SDS, 10% w/v dextran sulphate. The filter was washed in 2 x SSC, 1% w/v SDS at 65°C for between 1 to 2 hours and then 0.2 x SSC, 1% w/v SDS at 65°C for between 0.5 to 1 hour. The filter was exposed to Kodak XAR film with an intensifying screen at -70°C for 22 hours.

- 123 -

The T1 probe hybridized with transcripts of expected sizes (see Table 20) in RNA of transgenic plants that had been transformed with constructs carrying the T1 or A8 clones (lanes 1, 2, 5, 6, 7, 8, 13, 16 and 17) (eg. pCGP2772, pCGP2765, pCGP3259, pCGP2785, pCGP3258, pCGP3261) (Figure 41A) (Table 20). Under the conditions used, no hybridizing transcript was detected by Northern analysis of total RNA isolated from non transgenic control plants (lanes 9 and 10) or transgenic plants transformed with non-T1 carrying constructs (lanes 3, 4, 11, 12, 14 and 15) (e.g. pCGP960 (GUS), pBIN35Smgfp4, pCGP3260 (ER:mGFP4).

10

5

The <sup>32</sup>P-labelled T1 DNA probe was then stripped from the RNA blot by soaking the membrane in 0.1% SDS at 100°C and incubating it in a 65°C oven for 30 minutes with a final incubation step at room temperature for around 30 minutes.

15 The RNA blot was then probed with <sup>32</sup>P-labelled fragments of a ~0.8 kb *Hin*dIII fragment from pCGP1651 (SuRB) (10<sup>8</sup> cpm/μg, 2 x 10<sup>6</sup> cpm/mL). Prehybridization and hybridization were carried out as described above. The plasmid pCGP1651 contains a 0.8 kb *Hin*dIII fragment from the SuRB coding region contained in the binary plasmid vector pWTT2132 (DNAP).

20

25

The SuRB probe hybridized with a 2.2 kb transcript in transgenic plants that had been transformed with the constructs carrying the CaMV 35S: SuRB transgene (Figure 41 B) (lanes 1 to 8, 13 to 17) (eg. pCGP2772, pCGP2765, pCGP3259, pCGP2785, pCGP3258, pCGP3261) (Table 20). Under the conditions used, no hybridizing transcript was detected by Northern analysis of total RNA isolated from non transgenic control plants (lanes 9 and 10) or transgenic plants transformed with non-SuRB constructs (lanes 11 and 12) (e.g. pBIN35Smgfp4ER).

- 124 -

# Detection of colored proteins in transgenic Arabidopsis

Polyclonal rabbit antibodies to T1 protein

5 T1 protein was extracted from cultures of *E. coli* harbouring pCGP2921 (Figure 10) as described previously in Example 6.

Polyclonal rabbit antibodies against the T1 protein were produced by Institute of Medical and Veterinary Sciences, Veterinary Services Division, 101 Blacks Rd, Gilles Plains, South Australia 5086, Australia. An amount of 300 µg of T1 protein (with Freunds complete adjuvent) was initially administered. Serial doses of 300 µg T1 protein (with Freunds incomplete adjuvent) were subsequently administered 22 days and 36 days after the initial dose. Antibodies collected in the first bleed (which was taken at 45 days after the initial dose) were used to probe Western blots in the first instance.

15

20

30

10

#### Protein extraction from plants

Leaf material (20 -120 mg) was collected from *Arabidopsis* plants, snap frozen in liquid nitrogen and then ground to a fine powder using a mortar and pestle. An equal volume (w/v) of extraction buffer (100 mM Na<sub>2</sub>PO<sub>4</sub>pH 6.8, 150 mM NaCl, 10 mM EDTA, 10 mM DTT, 0.3 % Tween 20, 0.05 % Triton X) was then added to the fine powder and the mixture was further ground using the mortar and pestle. The resultant slurry was centrifuged at 10 000 rpm for 10 min and the supernatant was collected.

# 25 Western blot analysis of proteins extracted from transgenic Arabidopsis

Aliquots (8  $\mu$ L) of the protein extracts were mixed with 2  $\mu$ L of 5 x SDS loading buffer 10% v/v glycerol, 3% w/v SDS, 3%  $\beta$ -mercaptoethanol, 0.025% w/v bromophenol blue) electrophoresed through precast SDS PAGE gels (12% w/v resolving, 4% w/v stacking gel) (Ready Gels, Biorad) at 100 V for 1h 15 min in a Min-Protean System (Bio-Rad) using conditions as described previously in Example 6. The proteins were then transferred

- 125 -

to Immun-Blot PVDF membrane (Bio-Rad) using a Mini Trans-Blot Electrophoretic Transfer Cell (Bio-Rad) in Towbin buffer (25 mM Tris, 20 % methanol, 192 mM glycine) at 100 V for 1 h. PVDF membranes were incubated in blocking buffer (5 % non-fat dry milk, 0.2 % Tween-20, 75 mM NaPi pH 7.4, 68 mM NaCl) at room temperature for 1 h. Membranes were then further incubated with Rabbit anti-T1 antibody (diluted 1/200 in blocking buffer) for 2 h at room temperature then washed twice for 5 min in wash buffer (0.2 % Tween, NaPi pH 7.4, 68 mM NaCl). The membranes were finally incubated with goat anti-rabbit-IgG-alkaline phosphatase congugate (Bio-Rad) (diluted 1/300 in blocking buffer) for 1 h at room temperature followed by 4 washes for 10 min each in wash buffer. Colorimetric detection was carried out with Western Blue Stabilized Substrate for Alkaline Phosphatase (Promega).

5

10

15

20

25

The polyclonal T1 antibody detected a protein band running at the same position as T1 protein extracted from *E.coli* cultures harbouring pCGP2921 in extracts from *Arabidopsis*/2772 event 1.2, *Arabidopsis*/3259 event 1.5. The same T1 protein band was not detected in extracts from the non-transgenic controls.

The protein content in a 2  $\mu$ L sample of the protein extracts was estimated using a Bio-Rad Protein Assay as per the manufacturers instructions (Microassay Procedure). The absorbance of each extract at 595 nm was compared with BSA standard curves (0 - 10  $\mu$ g/mL) to estimate protein concentrations.

Samples of protein extract and a dilution series of known amounts of purified His-tagged colored protein (T1) were electrophoresed through SDS PAGE gels as described previously. The proteins were transferred to PVDF membranes (as described above) and probed with rabbit anti-T1 antibodies. The amounts of T1 colored protein in the protein extracts was estimated by comparison with the purified His-tagged colored protein dilution series. This allowed an estimation of expression of colored protein in *Arabidopsis* leaf as a percentage of total soluble protein (Table 21).

- 126 -

#### **EXAMPLE 15**

# Expression of colored proteins in Petunia

#### Transformation of petunia

5

Petunia hybrida cultivar Mitchell produces white flowers. Mitchell was transformed with the binary constructs pCGP2772 (Figure 24), pCGP2765 (Figure 21), pCGP3259 (Figure 36) pCGP2785 (Figure 33) and pCGP2926 (Figure 44) via Agrobacterium-mediated transformation as described in Example 1.

10

15

20

#### Northern analysis

Flowers from a random selection of events transformed with the T-DNAs of pCGP2772 and pCGP2765 were analysed for the presence of transcripts of the introduced T1 or A8 colored protein. Total RNA was isolated using a Plant RNAeasy kit (Qiagen) following procedures recommended by the manufacturer. Northern analysis was performed as described above for analysis of the *Arabidopsis* transgenic plants.

The T1 probe hybridized with transcripts of around 0.9 kb in petal RNA of transgenic Mitchell plants that had been transformed with constructs carrying the T1 or A8 clones (Figure 40A) (pCGP2772 (lanes 7 to 12) and pCGP2765 (lanes 1 to 6), respectively). Under the conditions used no hybridising transcript was detected in RNA isolated from petals of a non transgenic control (data not shown).

The SuRB probe hybridized with a 2.2 kb transcript in transgenic plants that had been transformed with the constructs carrying the CaMV 35S: SuRB transgene (Figure 40B).

Under the conditions used no hybridizing transcript was detected in RNA isolated from petals of a non transgenic control (data not shown).

30

- 127 -

## Western blot analysis of proteins extracted from transgenic Petunia

5

20

25

30

Proteins were extracted from leaf and flower material (petal tube, petal limb, anthers, pistil, stigma and style) (100 - 300 mg) of transgenic and non-transgenic *P. hybrida* cv, Mitchell plants as described for *Arabidopsis*.

Western blot analysis of these protein extracts was performed as described for Arabidopsis.

- The polyclonal T1 antibody detected a protein band running at the same position as T1 protein extracted from *E.coli* cultures harbouring pCGP2921 in extracts from *Petunia* accession 24534 (pCGP2765) and *Petunia* accession 24444 (pCGP2772). The same T1 protein band was not detected in the non-transgenic controls.
- An estimation of expression of colored protein in *Petunia* leaf and petal as a percentage of total soluble protein was made as described above for *Arabidopsis* extracts (Table 22).

The T1 protein was produced in *Arabidopsis* leaf (Example 14) and *Petunia* leaf and flower tissue (Example 15). It is expected that an increase in protein accumulation will produce stronger colours in flower and leaf tissue. The first generation of transformed plants are selfed to give homozygous second generation transformants with higher T1 protein or other CFM accumulation and stronger colour.

Alternatively, different transgenic events are crossed to produce second generation transformants with higher T1 protein or other CFM accumulation and stronger colour. Methods envisaged to increase total T1 protein or other CFM accumulating in transformed plants include targeting T1 or other CFM to the chloroplast using a chloroplast transit peptide such as that from the small subunit of ribulose-bisphosphate from tobacco (see Example 11 or Table 17). These chloroplast transit peptides will facilitate the movement and accumulation of CFMs into chloroplasts which are abundant in leaves and chromoplasts which are abundant in flowers petals. Another method envisaged to produce

- 128 -

higher levels of CFMs in plant tissues is the use of chloroplast/plastid transformation techniques which have been used in the past to generate plants expressing recombinant proteins at levels of up to 46 % of total soluble protein (De Cosa et al., Nat. Biotechnol. 19, 71-74, 2001; Daniell et al., Trends in Plant Sci. 7: 84-91, 2002, see Example 11, Table 18). It is also envisaged that the co-expression of a suitable chaperonin in conjunction with one or more CFMs allows the efficient folding and packaging of CFMs into stable structures which are accumulated in higher amounts than would normally be expected. It is also envisaged that producing a fusion of CFM with ubiquitin in plants will increase levels of accumulated CFMs in transgenic plants as has been demonstrated in yeast (Baker, Curr. Opinions in Biotech, 7: 541-546, 1996 and references within). It is also envisaged that targeting T1 or other CFM to the endoplasmic reticulum (see Example 11) will increase the levels of accumulated recombinant protein in plant tissues (Haseloff et al., 1997, supra).

# Detection of correctly folded CFMs in plant extracts.

15

20

25

10

5

CFMs that are folded correctly in heterologous systems (such as when expressed in flowers or other plant tissues) are expected to retain characteristic absorbance and corresponding colour (see Example 13). The level of CFM production or accumulation may initially be too low for significant color change in plant tissue. A method for detecting low levels of correctly folded CFMs in plant extracts is described for leaf material from *Petunia* transformed with pCGP2772 and pCGP2765, however, this method can be used with other plant tissues such as but not limited to *Petunia* or rose or gerbera.

Total soluble proteins were extracted from transgenic leaves of Mitchell/pCGP2772 and Mitchell/pCGP2765) (see Example 15). These samples were frozen in liquid nitrogen and ground using a mortar and pestle. An equal volume (w/v) of extraction buffer (100 mM NaPO4 pH 6.8, 150 mM NaCl, 10 mM EDTA, 10 mM DTT, 0.3 % Tween 20, 0.05 % Triton X) was added to the sample and further ground. The resultant slurry was centrifuged at 10 000 rpm for 10 min and the supernatant collected.

- 129 -

The extracts were used undiluted or diluted 1:2 in water and their absorbance characteristics determined between 400 nm and 700 nm using a Varian Cary 50 Bio UV-Visible Spectrophotometer. The absorbance spectra were compared to those of extracts of non-transgenic control tissue and non-transgenic control tissue spiked with either T1 or T3 His-tagged purified protein (see Example 8). Detectable color was observed through the detection of peaks at approximately 580-590 nm in the extracts from transgenic plant tissue that were not evident in non-transgenic control tissue.

Methods envisaged to increase protein levels are as described above or by Bailey-Serres and Gallie (American Society of Plant Physiologists, Look beyond transcription, UCLA, USA, 1998) or by modification of mRNA sequence to optimize 5' and 3' untranslated sequences thereby improving message stability and/or translation efficiency, optimisation of codon usage in the introduced gene to more closely match that found in highly expressed genes (that is genes which give rise to high levels or encoded protein synthesis) in particular those of target crops, augmentation of protein stability via the attachment for example of stabilising sequences such as ubiquitin, changes to specific N-terminal amino acid residues to promote altered aggregation of monomeric forms of the protein, more effective targeting of the synthesized polypeptide to intracellular organelles or compartments, duplication and there for amplification of introduced genes leading to increased levels of protein biosynthesis for example using 'Gene Amplification Technology' (Borisjuk et al., Nature Biotechnology 18: 1303-1306, 2000).

#### **EXAMPLE 16**

## Expression of colored proteins in other plants

25

30

20

5

10

15

The horticultural industry relies on the production of novel traits such as new colors, fragrances, productivity and disease resistance. Introduction of colored protein sequences (via genetic engineering) into commercially important plant lines such as, for example, but not limited to roses, carnations and gerberas provides a means to produce novel colors in flowers or plants that lack such colors.

- 130 -

Introduction of colored protein genes into roses is achieved using methods such as those described, for example, in International Patent Application Number PCT/US91/04412, or by Robinson and Firoozabady (*Scientia Horticulturae*, 55: 83-99, 1993), Rout et al. (*Scientia Horticulturae*, 81: 201-238, 1999) or Marchant et al. (*Molecular Breeding 4*: 187-194, 1998) or by any other method well known in the art.

Introduction of colored protein genes into carnations is achieved using methods such as those described, for example, in International Patent Application Number PCT/US92/02612 or by Lu et al. (Bio/Technology 9: 864-868, 1991), Robinson and Firoozabady (1993, supra) or by any other method known in the art.

Introduction of colored protein genes into carnations is achieved using methods such as those described, for example, by Robinson and Firoozabady (1993, supra).

The cotton industry relies on the production of dyed cotton, using dyes that can have concomitant detrimental effects on the environment. Introduction of colored protein sequences (via genetic engineering) into commercially important cotton lines, or other plant lines that allow for production of fabrics (such as, but not limited to, hemp), and also relies on use of colored dyes to dye said fabrics, is achieved using methods such as those described, for example, in an International Patent Application having Publication Number WO 00/77230.

#### **EXAMPLE 17**

#### Generation of transformed animals

25

30

5

10

The use of the CFMs of the present invention are employed to produce transgenic animals which exhibit novel color, for example, sheep with blue or red colored fleece, cows with red colored hide *inter alia*. The transgenic animals of the present invention can be produced by any number of method know in the art. Such as, but not limited to transgenic animals are produced by any number of methods, for example, microinjection of constructs

- 131 -

comprising a CFM nucleotide sequence into the pronucleus of a fertilized ovum, or injection of embryonic stem (ES) cells into embryos.

### Microinjection

5

10

20

25

Following fertilization a single celled embryo is removed from the animal (e.g. sheep, cow, pig, goat). Micromanipulators on a specially equipped microscope are used to grasp each embryo. A glass pipette drawn to a fine point immobilizes the embryo on one side. On the opposite side, a construct containing a CFM nucleotide sequence is injected into the embryo's pronucleus with a second finely drawn injection needle. Following the injection, the embryos are transferred back into the hormonally prepared or pseudopregnant recipient females or foster mothers. The recipients follow normal pregnancy and deliver full-term young.

# 15 Injection of embryonic stem cells

ES cells are isolated from the inner cell mass of blastocyst-stage embryos (about 7 days postfertilization), ES cells are grown in the lab for many generations to produce an unlimited number of identical cells capable of developing into fully formed adults. These ES cells are altered genetically by injection of a construct containing a CFM nucleotide sequence.

Transgenic individuals are produced by microinjection of embryonic stem (ES) cells containing the CFM construct into embryos to produce "hybrid" embryos of two or more distinct cell types. Following the injection, the embryos are transferred back into the hormonally prepared or pseudopregnant recipient females or foster mothers. The recipients follow normal pregnancy and deliver full-term young.

- 132 -

#### **EXAMPLE 18**

## Generation of a far red fluorescent monomeric protein

#### Cloning and expression

5

cDNA encoding the colored protein Rtms-5 (SEQ ID NO:166) was isolated from *Montipora efflorescens* (Scleractina Acropodiae). Under daylight illumination, *Montipora efflorescens* was a purply-red colour, but fluoresced yellow under blue illumination and red under green illumination.

10

15

30

To further characterise the protein, the cDNA was tagged with hexahistidine at its C-terminus and expressed at high levels in *Escherichia coli*. For expression in bacteria, the nucleotide sequence encoding Rtms-5.pep (SEQ ID NO:166) was retrieved from pGEM-T cloning vector (Promega) using forward oligonucleotide primers consisting of the NotI restriction binding site, a ribosomal binding site, a spacer and 15 bases encoding the N-terminus of the protein (MSV-RBS, SEQ ID NO:213; SVIAK-RBS, SEQ ID NO:214) and a reverse oligonucleotide primer encoding H6-tag (POC220-H6, SEQ ID NO:215).

#### SEQ ID NO:213 MSV-RBS

20 GGC TCT AGA AAG GAG ATA TAC AAG TGT GAT CGC TAC ACA AAT GA

## SEQ ID NO:214 SVIAK-RBS

GGC TCT AGA AAG GAG ATA TAC AAT GTC CGT TAT CGC TAA ACA GAT

# 25 SEQ ID NO:215 POC220-H6

GGC AAG CTT TCA GTG GTG GTG GTG GTG GGC GAC CAC AGG TTT GCG TG

PCR product was gel purified and diluted (x10) prior to cloning into pCRII-TOPO (Invitrogen) and transforming into Top 10 cells (Invitogen). Cells were induced with 0.5mM IPTG, and protein was purified on Ni-columns (Pro-Bond, Invitrogen) eluting with 50mM, 200 mM, 350 mM and 500 mM Imidazole in PBS pH 6.0, prior to overnight dialysis against 50 mM Potassium Phosphate pH 6.65.

- 133 -

#### Fluorescence charcteristics of Rtms-5

5

10

15

20

30

E. coli colonies were blue in colour in daylight, and weakly red fluorescent when excited with light of wavelength 595 nm.

An alignment of the amino acid sequence of Rtms-5 (SEQ ID NO:166) with other fluorescent proteins was constructed (Table 19). Rtms-5 (SEQ ID NO:166) contains the key amino acids (Tyr-66 and Gly-67) that correspond to those that form the fluorophore in other well-characterised proteins, dsRed583 (also known herein as drFP583, SEQ ID NO:221) and GFP (SEQ ID NO:222). Overall, 67% and 20% of the Rtms-5 (SEQ ID NO:166) sequence is identical to dsRed583 (SEQ ID NO:221) and GFP (SEQ ID NO:222), respectively. The protein shares a high degree of identity with a number of chromoproteins recently isolated from the *Anthozoa* species (Gurskaya et al., FEBS Lett. 507: 16-20, 2001).

The absorption and excitation emission spectra were measured for the purified "wild-type" Rtms-5 (SEQ ID NO:166). The protein displays a major absorption peak at 592 nm, with an extinction that is highly variable ( $\varepsilon_{592} = 53,000 \,\mathrm{M}^{-1} \,\mathrm{cm}^{-1}$ –111,000  $\mathrm{M}^{-1} \,\mathrm{cm}^{-1}$ ) and a shoulder peak at 454 nm (Figure 42. The variability in the extinction coefficient is similar to that observed for drFP583 (SEQ ID NO:221) and, similarly, it is dependant on the state of maturity, as well as the conditions under which the protein is expressed (Baird *et al.*, 2000, *supra*).

#### 25 Site directed mutagenesis

Rtms-5 (SEQ ID NO:166) was only weakly fluorescent. To enhance this, site-directed mutagenesis was carried out. The alignment of the Rtms-5 sequence (SEQ ID NO:166) with other sequences (Table 19) indicated that position 142 was occupied by the residue histine. A variant Rtms-5-H142S, containing the substitution H142S, was engineered by mutagenesis of pCRII-TOPO::RTms5 to produce pCRII-TOPO::RTms5-H142S. This

- 134 -

single substitution increased the quantum yield of far-red fluorescence by 170-fold to a quantum yield of less than 0.02. Minor effects on the excitation and emission spectra and the absorption spectra were observed (4 nm shift towards the blue end of the spectrum, refer to Figure 42A,B,C).

5

10

15

20

25

#### Analysis of oligomeric structure

dsRed583 (SEQ ID NO:221) is known to be an obligate tetramer. The formation of oligomers by fluorescent proteins can present a serious problem when expressed fused to other proteins of interest. Consequently, it was important to establish the degree of oligomerisation of Rtms-5 (SEQ ID NO:166). The protein has a predicted size of 25,820 Da (with H6). When subjected to SDS-PAGE under reducing conditions, purified Rtms-5 (SEQ ID NO:166) migrated with an M<sub>r</sub> of 26,900. However, under non-reducing conditions the majority of the protein migrated with an M<sub>r</sub> of 114,000. These results indicated that native Rtms-5 (SEQ ID NO:166) was predominantly a tetramer.

#### Further site directed mutagenesis and analysis of structure

A second round of site-directed mutagenesis was carried out, to mutagenise CRII-TOPO::RTms5-H142S to produce the variant pCRII-TOP-RTms5-H142S-F158H (pCRII::Rtms-5v). This colored peptide contained the additional substitutions F158H and R145H, and is designated Rtms-5v (SEQ ID NO:216).

Rtms-5v (SEQ ID NO:216) was expressed in *E. coli* and the purified six His-tagged protein was subjected to analytical ultracentrifugation. The results indicated that the mutagenised variant sedimented predominantly as a monomer (82%, 30,700 Da) with the remaining proportion sedimenting as a dimer (18%, 50,800 Da). This colored protein fluoresced in the far-red range (see Figure 42C), and can be used effectively in yeast cells and mammalian cells.

- 135 -

# Effect of site directed mutagenesis of other colored proteins

Site directed mutagenesis of residue H or N 142 to S, in other colored protein sequences, also leads to the generation of far-red fluorescence. Examples of the excitation and emission spectra for two other colored proteins, Aams-4 (SEQ ID NO:90)-H142S, and Rtms-1 (SEQ ID NO:162)-N142S are shown in Figure 43.

#### **EXAMPLE 19**

# Expression in yeast, mammals and as a fusion protein

10

5

The subject inventors sought to demonstrate that the instant CFMs can be expressed in yeast and mammalian cells and can be used as fusion proteins for genetic marking of cells.

## (a) Expression in yeast

15

20

25

30

For expression in yeast cells a BamHI/Not1 DNA cassette encoding dsRed or YGFP3 (an enhanced variant for expression in yeast) or a BgIII/Not1 cassette encoding the novel fluorescent protein, Rtms-5v (SEQ ID NO:216), were retrieved using the pair of oligonucleotide primers RFPUP1 (SEQ ID NO:234), /RFPDO1 (SEQ ID NO:235), YGFP3UP (SEQ ID NO:232), /YGFP3DO (SEQ ID NO:233), or MSVIATUP (SEQ ID NO:236)/COFPDO (SEQ ID NO:237), respectively, using as templates the vectors pYGFP3 (Cormack et al., Microbiology 143: 303-11, 1977), pDsRed-1 [Clontech Industries] or cDNA for pCRII-TOPO::RTms-5v. In the case of YGFP3UP, the Not1 site was retrieved after digesting the PCR product from pGEM-T (Promega). The PCR product was cloned into the BamHI/Not1 site of the multi-copy yeast expression vector pAS1NB to produce pAS1NB::dsRedL, pAS1NB::YEGFP3L or pAS1NB::Rtms-5v from which the DNA cassette encoding wild-type GFP had been removed but retaining the multiple cloning sites of that vector and linker sequence of that vector [Prescott et al., FEBS Letts. 411: 97-101, 1997]. pASN1B is a derivative of pAS1N (Prescott et al., 1997, supra) in which a BamH1 restriction site has been removed from the PGK promoter region. This

- 136 -

series of vectors allows the expression of fluorescent proteins not fused to a partner protein and provides.

SEQ ID NO:232 YGFP3UP

5

5'- GGATCCATCGCCACCATGTCTAAAGGTGAAGAATTATTCACTGG

SEQ ID NO:233 YGFP3DO

10 5'- CAGCTGTTATTTGTACAATTCATCCATACCATGG

SEQ ID NO:234 RFPUP1

5'- CGGGATCCATCGCCACCATGAGGTCTTCCAAGAATGTTATC

15

SEQ ID NO:235 RFPDO1

- 5'- GAGGATCCGCGCCCCTAAAGGAACAGATGG
- 20 SEQ ID NO:236 MSVIATUP
  - 5'- GAAGATCTAAAACAATGAGTGTGATCGCTACACAAATG

SEQ ID NO:237 COFPDO

25

- 5'- TATCAAATCGCCGGCGTCAGGCGACCACAGGTTTG
- (b) Expression as a fusion protein
- 30 Two DNA cassettes encompassing segments of the yeast genes ATP4 and ATP7 for subunit b and d of ATP synthase, respectively, were recovered by PCR from YRD15

- 137 -

genomic DNA using the oligonucleotide primer pairs ATP4PROMUP2 (SEQ ID NO:238)/ATP4DO2 (SEQ ID NO:239), or ATP7TUP (SEQ ID NO:240)/ATP3TDO (SEQ ID NO:241), respectively. The first, ATP4PO, encompasses the open reading frame for ATP4 and 500 bp of sequence upstream of the initiation codon flanked by Bg/III and BamHI restriction sites at the 5' and 3', respectively. The BamH1 restriction site allows for an in frame-fusion between the C-terminus of subunit b and each of the three fluorescent protein cassettes. The second, ATP7T, encompasses the transcription terminator cassette representing the terminator region of the ATP7 gene flanked at the 5' and 3' ends by restriction sites for NotI and SacII, respectively. These restriction sites were obtained on cloning the PCR product into GEM-T. The ATP4PO & ATP7T DNA cassettes were cloned sequentially into the BamHI and NotI/SacII sites, respectively of the yeast expression vector pRS413 to produce the expression vector construction denoted pRS413::ATP4PO:ATP7T. A Bg/IIIHI/NotI DNA fragment encoding YGFP3L was excised pAS1NB::YEGFP3L and then cloned into the BglII/NotI site pRS413::ATP4PO:ATP7T to produce a vector (pRS306::ATP4PO:YEGFP3L:ATP7T) encoding subunit b fused to YEGFP3 with a polypeptide linker of 25 amino acids. A vector (pRS413::ATP4PO:RTms-5:ATP7T or pRS413::ATP4PO:dsRed:ATP7T) encoding subunit b fused to RTms-5B or dsRed with a polypeptide linker of 27 amino acids was derived from pRS306::ATP4PO:YGFP3L:ATP7T by replacing the BamHI/Not1 fragment encoding YEGFP3 with an equivalent fragment encoding Rtms-5v or dsRed.

SEQ ID NO:238 ATP4PROMUP2

5'- AGATCTGTGTTGTGACGCAACTCC

**?**5

5

10

15

20

SEQ ID NO:239 ATP4DO2

5'- GTGATCAGCGGATCCCTTCAATTTAGAAAGCAATTGTTC

- 138 -

#### SEQ ID NO:240 ATP7TUP

# 5'- CCTCTATATATTACGCACCATATTC

#### 5 SEQ ID NO:241 ATP7TDO

10

15

20

25

30

#### 5'- ATACGTGACGACATTGGTAGTC

# (c) Results were visualised using Clear Native Gels.

These were run essentially as described hereinaster. Briefly, 200 μg of mitochondrial protein was pelleted for 5 min at 100,000 g. Yeast mitochondria were isolated from spherophlasts (Law et al., Methods in Enzymol. 260: 122-163, 1995). The pellet was solubilized in buffer (40 μl) containing in dodceyl β-maltoside to isolate the monomer form or digitonin (20 g/g protein) to isolate the dimer form and incubated on ice for 20 min and centrifuged 100,000 g for 30 min. Supernatants (30 μl) were loaded into wells of 4-16% gradient gels (13 cm x 10 cm x 0.075 cm). After running and while still between the glass plates, gels were imaged for fluorescence using a Perkin-Elmer multi-wavelength imager in 'edge-illumination mode' using appropriate filters for excitation (GFP, 480±20 nm; dsRed and Rtms-5v, 540±25 nm) and emission (GFP, 535±20 nm; dsRed, 590±35 nm; Rtms-5v, 620±30 nm).

DNA cassettes encoding subunit b fused to the N-terminus of each of the three fluorescent proteins were expressed in a yeast strain lacking expression of endogenous subunit b. The ATP synthase in each of these strains was established to be assembled and functional as cells of each strain were able to grow using the non-fermentable substrate ethanol as carbon source. Yeast cells lacking endogenous subunit b do not assemble functional mtATPase and cannot grow using ethanol as the sole carbon source. Yeast cells of each strain expressing the individual fusion proteins were visualized using fluorescence microscopy. For cells of each strain the distribution of fluorescence in the cell was similar and consistent with localisation to the mitochondrion.

- 139 -

Mitochondria were isolated from cells of each of the strains and, after extraction, ATP synthase complexes were subjected to analysis by clear native gel electrophoresis (CNGE). ATP synthase isolated from yeast is a large membrane bound complex (~800 kDa for the monomeric form) made up of 20 different types of subunits some of which are present in the complex as more than one copy. The complex can be isolated as a monomer or a dimer depending on the detergent, dodceyl \beta-maltoside or digitonin, respectively, used to extract the complex from mitochondrial membranes. Subunit b is present in a single copy in the monomer. ATP synthase in this experiment was extracted from each preparation of mitochondria under conditions that favour the isolation of the monomer. Subunit b is present in a single copy in the monomer. Samples were subjected to analysis by CNGE and the gel imaged for fluorescence using conditions of illumination and light detection specific for each fluorescent protein (Figure 47). A single fluorescent band corresponding to the position of assembled monomeric ATP synthase was observed for complexes containing the b-GFP fusion protein (Figure 47, lane 1). The position of GFP not fused to another protein is shown (Figure 47, lane 4). A single fluorescent band was seen for complexes containing the fusion protein b-Rtms-5v (Figure 47, lane 2). However, multiple bands were observed for samples containing b-dsRed Figure 47, lane 3). It is possible that, in order of decreasing mobility, each fluorescent band corresponds to a monomer, dimer, trimer and tetramer.

# (d) For expression in mammals

5

10

15

20

25

For expression in mammalian cells, a *Smal/Not*I fragment encoding Rtms-5v (SEQ ID NO:216) was excised from pAS1NB::RTms-5v and cloned into the expression site of the mammalian expression vector pCI-Neo (Promega Corporation, Madison USA). This vector allows the expression of Rtms-5v not fused to a partner protein.

A major benefit of fluorescent protein technology is the ability to simultaneously monitor using spectrally distinct variants more than one event in the living cell. The spectral properties of Rtms-5v suggest that should be feasible to image both dsRed and Rtms-5v

- 140 -

expressed in the same cell. This would allow Rtms-5 to be used in combination with dsRed rather than substitute for dsRed. The emission maxima for dsRed and Rtms-5v are separated by 50 nm. We tested the possibility of imaging dsRed, RTms-5v and EGFP expressed in the same cell. Three individual DNA cassettes were constructed encoding dsRed fused at its N-terminus to the 16 amino acid mitochondrial targeting sequence of human 3-oxoacyl-CoA thiolase, EGFP fused to the C-terminus of Rab6 and Rtms-5v not fused to any other protein. Cells were imaged using a Zeiss 510 Meta confocal laser scanning microscopy (Zeiss). The distribution of fluorescence arising from each of the Rtms-5v, dsRed and EGFP fusions was consistent with the locations expected (cytosol/nucleus, mitochondrion and golgi, respectively). These results show that Rtms-5v is able to fluorescently label other compartments of the cell such as the mitochondrion in addition to the cytoplasm. The position of a non-transfected and, therefore, non-fluorescent cell is shown in the transmitted light image by the white arrow Rtms-5v showed no evidence of aggregation. Similar results were observed for the expression of Rtms-5v not targeted in yeast cells. Multiple fluorescent proteins are commonly (eg. GFP, dsRed, CFP) imaged in the same cell.

# **EXAMPLE 20**

# Additional color proteins from coral

20

15

10

The inventors sought additional color proteins from two corals, *Montipora efforescens* and *Pavona decussaca*.

#### (a) Montipora efforescens

25

30

Standard purification techniques (Dove et al., 2001, supra) were adopted for the purification of a red fluorescent protein from phosphate buffer extract of M. efforescens. A protein was purified using gel filtration and subject to N-terminal amino acid sequencing. A polymorphism was identified, comprising F and R residues. The N-terminal amino acid sequences are represented as follows:

- 141 -

SPPDY TLEFP KKXVA

**SEQ ID NO:242** 

SPPDY TLERP KKGVA

**SEQ ID NO:243** 

The polymorphism is indicated in bold larger type.

5

# (b) Pavona decussaca

Similar techniques as those described in (a), above, were used to identify and purify a green fluorescent protein from *P. decussaca*. Gel electrophoresis showed that the proteins ran as two bands and N-terminal amino acid sequencing identified polymorphic variants, shown in bold larger type, below:

Top band:

(D)SS(P)E SYLKN GIAEE MKTDV MEGI

**SEQ ID NO:244** 

15

25

10

Lower band:

SYLPN GIAEE MKTDL MEGIV NG

**SEQ ID NO:245** 

SLYQN GIAEE MKTDL MEGIV NG

**SEQ ID NO:246** 

The protein fraction was generating these N-terminal sequences had absorbed maximally at 440 nm with maximal excitation at 440 nm and emission at 488 nm.

Oligonucleotide probes were designed in both forward and reverse directions for PCR amplification from a ZAP express cDNA library of *Acropora millepora* (Scleractinian coral). The oligonucleotide primers used were as follows:

#### **Forward**

**MEGIVNG-A** 

ATG GAA GGG ATA GTC GAT GG

**SEQ ID NO:247** 

30 MEGIVNG-T

ATG GAA GGG ATT GTC GAT GG

**SEQ ID NO:248** 

- 142 -

MEGIVNG-C ATG GAA GGG ATC GTC GAT GG SEQ ID NO:249

Reverse

10

15

20

5 REV-MEG-T CCT CGA CAA TCC CTT CCA T SEQ ID NO:250

REV-MEG-C CCT CGA CGA TCC CTT CCA T SEQ ID NO:251

DNA was amplified and separated using gel electrophoresis. Bands were purified and cloned into pCRII-TOPO and transfected into TOP 10 cells (Invitrogen). Plasmids were then purified and subjected to nucleotide sequencing. The complete sequence is shown in Table 23.

In this experiment, therefore, a protein identified from *P. decussaca* was used to identify a clone from *Acropora millepora*.

Those skilled in the art will appreciate that the invention described herein is susceptible to variations and modifications other than those specifically described. It is to be understood that the invention includes all such variations and modifications. The invention also includes all of the steps, features, compositions and compounds referred to or indicated in this specification, individually or collectively, and any and all combinations of any two or more of said steps or features.

%identity(length) *		95.0(220) - 97.3(221)								43.5 <sup>(225)</sup> & 43.5 <sup>(225)</sup>			52.7 <sup>(216)</sup>	46.2 <sup>(227)</sup> & 47.2 <sup>(223)</sup>	93.2 <sup>(220)</sup> _ 100 <sup>(255)</sup>	44.1(231) & 45.5(236)	57.8 <sup>(225)</sup> & 62.5 <sup>(224)</sup>	
Sequence information	Prasher et al. (1992) Gene 111:229-33; Chalfie et al. (1994) Science 263:802-5; Inouye & Tsuji (1994) FEBS Lett. 341:277-80.	Invention								WO 99/49019			Matz et al., Nature Biotechnology 17: 969-973, 1999	Matz et al., Nature Biotechnology 17: 969-973; 1999; Lukyanov et al., JBC 275: 25879-25882, 2000	WO 00/46233; Dove <i>et al., Coral</i> Reefs 19:197-204, 2001; Invention.	Matz et al., Nature Biotechnology 17: 969-973, 1999	Matz et al., Nature Biotechnology 17: 969-973, 1999	
Initial studies	Shimomura et al., J. Cell.  Comp. Physiol. 59: 223-239, 1962; Morise et al., Biochemistry 13:2656-2662, 1974; Morin & Hastings, J. Cell Physiol. 77: 313-318, 1971									Morin & Hastings, J. Cell Physiol, 77: 313-318, 1977					Kawaguti, Paloa Trop. Boil. Sm. Stud. 2: 617-673, 1994; Dove et al. Biol. Bull. 189: 288- 297, 1995			
GFP	yes	yes						·		yes			yes	yes	yes	yes	yes	
Order	Hydroida	Milliporina	Stylasterine	Trachylina	Siphonophora	Chondrophora	Actinulida	Gorgonacea	Telestacea	Pennatulacea (see pens & sea pansies)	Alcyonacea	Helioporacea	Stolonifera	Actiniaria (sea anemones)	Scleractinia (true or stony corals)	Zoanthidea	Corallimorpharia (coral-like anemones)	Antipatharia
SubClass						•		Octocorallia (=Alcyonaria)						Hexacorallia (=Zoantharia)				Ceriantipatharia
· Class	Hydrozoa							Anthozoa							·			

	1	44	
_		n	_

* (q)Baia							-
%identity(	131107				•		
Sequence information							
Initial studies				***	_		
GFP			1	3			
Order	Ceriantharia		Stauromedusae	Coronatae	Semaeostomae	Rhizostaomae	
SubClass							
Class		Cubozoa	Scyphozoa (jellyfish)				

Best fit in relation to Aams2-pep (SEQ ID NO:88) over 220-238 amino acids as indicated in length

TABLE 3 Fluorescent properties

Excitation region	Exciter filter	Dichroic mirror	Additional barrier filter
Ultra-violet	UG-1	DM400 + L420	L435
Violet	BP 405	DM455 +Y455	Y475
Blue	BP 490	DM500 + O515	O515
Green	BP 545	DM580 + O590	R610

TABLE 4 Class: Anthozoa; Order: Scleractinia

Family	Genus, Species	Color morph	Fluorescent properties
Pocilloporidae	Pocillopora damicormis	Pink	Faintgreen fluorescence under blue light
	Pocillopora damicormis	Green	Fluoresce blue-green, green, green and red under UV, violet, blue and green light, respectively
Acroporidae	Acropora aspera	Blue tipped	Tentacles and calices fluoresce violet, blue, green, and faint red under UV, violet, blue and green light, respectively
	Acropora aspera	Blue light fluorescent	Fluoresces green under UV and violet light
	Acropora nobilis	Green	Calices and tentacles fluorescent violet, blue, green and red under UV, violet, blue and green light, respectively
	Montipora sp. (plating)	Red/red fluorescent	Yellow fluorescence under blue light, red fluorescence under green light
Poritidae	Porites murrayensis	Purple	Calices fluoresce faint green under violet and blue light
Agariciidae	Pavona decussaca	Green	Blue under UV light; green under violet light; and blue and moderate red under green light
Mussidae	Acanthasthastrea	Green	Calices and polyps fluorescent violet, blue, green and faint under UV, violet, blue and green light, respectively
Faviidae	Platygyra sp.	Green	Blue under UV and violet light; green under blue light
	Caulastrea sp.	Green	Blue under UV light; green under violet and blue light

- 146 -

TABLE 5 Class: Hydrozoa; Order: Milleporina

Genus	Color	Fluorescent properties
Millepora	Green	Blue under UV and violet light; green under blue light

N-terminus	Genus species	Name	Type	Amino acids within 5. A of fluorophore
sgiat	Acropora aspera	Aams-5.pep	1	QVLSPQCQYGSIFWRNSYEHENMERLQCE
sviat	Acropora aspera	Aams-2.pep	1	OVLSPOCOYGSIFWRNSYEHENMERLOCE
sviat	Acropora aspera	Aams-4.pep	1	QVLSPQCQYGSIFWRNSYEHENMERLQCE
sviat	Acropora aspera	Aams-6.pep	1	QVLSPQCQYGSIFWRNSYEHENMERLQCE
sviat	Acanthastrea sp.	Acams-2.pep	<b>*</b> 9	QVLSPQYQYGSIYWRNSYENENMERLQCE
sviat	Acanthastrea sp.	Acarns-3.pep	9	QVLSPQYQYGSIYWRNSYENENMERLQCE
sviat	Acanthastrea sp.	Acams-4.pep	3	QVLSPQYQYGSIYWRNSHENENMERLQCE
sviat	Acanthastrea sp.	Acams-5.pep	*	
sviat	Caulastrea sp.	Cems-F.pep	5	QVLSPQCQYGNIFWRNSYEHENMGRLQCE
sviat	Caulastrea sp.	Cems-G.pep	5	QVLSPQCQYGNIFWRNSYEHENMGRLQCE
sviat	Caulastrea sp.	Cems-H.pep	5	QVLSPQCQYGNIFWRNSYEHENMGRLQCE
sviat	Caulastrea sp.	Cems-I.pep	16*	QVLSPQCQYGNIFWRNSYEHENMERLQCE
sviat	Acropora nobilis	LGAms-5.pep	9	QVLSPQYQYGSIYWRNSYENENMERLQCE
sviat	Acropora nobilis	LGAms-6.pep	18*	QVLSPQYQYGSIFWRNSYENENMERLQCE
sviat	Millepora sp.	Mi68Dms.pep	1	QVLSPQCQYGSIFWRNSYEHENMERLQCE
sviat	Millepora sp.	Mims-A.pep	1	QVLSPQCQYGSIFWRNSYEHENMERLQCE
sviat	Millepora sp.	Mims-B.pep	1	QVLSPQCQYGSIFWRNSYEHENMERLQCE
sviat	Millepora sp.	Mims-C.pep	1	QVLSPQCQYGSIFWRNSYEHENMERLQCE
sviat	Pavona decussata	Pav5ms.pep	9	QVLSPQYQYGSIYWRNSYENENMERLQCE
sviat	Pavona decussata	Pavms-2.pep	<b>*</b> 9	QVLSPQYQYGSIYWRNSYENENMERLQCE
sviat	Pavona decussata	Pavms-3.pep	9	QVLSPQYQYGSIYWRNSYENENMERLQCE
sviat	Pavona decussata	Pavms-4.pep	11	QVLSPQYQYGSIYWGNSYENENMERLQCE
sviat	Porites murrayensis	PMms-A.pep	2	QVLSPQSQYGSIYWRNSYENENMERLQCE
sviat	Porites murrayensis	PMms-B.pep	2	QVLSPQSQYGSTYWRNSYENENMERLQCE
sviat	Porites murrayensis	PMms-C.pep	6	QVLSPQTQYGSIYWRNSYENGNMERLQCE
sviat	Porites murrayensis	PMms-D.pep	2	QVLSPQSQYGSIYWRNSYENENMERLQCE
sviat	Porites murrayensis	PMms-E.pep	2	QVLSPQSQYGSIYWRNSYENENMERLQCE
sviat	Pink Pocillopora	PPd57-1ms.pep	12	QVLSPQTQYGSIYWRNSYENENMERLQCE

	1	40	
-		4X	-

sviat         Pink Pocillopora         PPd57-2ms.pep         1         QVLSPQCQYGSIFWRNSYEHENMERLQCE           sviat         Pink Pocillopora         PPd57-3.pep         2         QVLSPQCQYGSIFWRNSYEHENMERLQCE           sviat         Platygyra sp.         PPms-1.pep         8         RVLSPQCQYGNIFWRNSYEHENMERLQCE           sviat         Platygyra sp.         PPms-2.pep         19*         QVLSPQCQYGNIFWRNSYEHENMGRLQCE           sviat         Platygyra sp.         PPms-1.pep         5         QVLSPQCQYGNIFWRNSYEHENMGRLQCE           sviat         Montipora sp.         RTms-1.pep         6         QVLSPQCQYGNIFWRNSYEHENMGRLQCE           sviat         Montipora sp.         RTms-5.pep         13         QVLSPQCQYGSIFWRNSYEHENMERLQCE           sviat         Montipora sp.         RTms-6.pep         6         QVLSPQCQYGSIFWRNSYEHENMERLQCE           svivit         Montipora sp.         RTms-6.pep         6         QVLSPQCQYGSIFWRNSYENENDERLQCE           svivit         Montipora sp.         RTms-2.pep         6         QVLSPQCYGSIFWRNSYENENDERLQCE           svivit         Montipora sp.         RTms-2.pep         6         QVLSPQCYGSIFWRNSYENENDERLACE	N-terminus	Genus species	Name	Type	Amino acids within 5. A of fluorophore
Pink Pocillopora         PPd57-3.pep         1           Pink Pocillopora         PPd57-4ms.pep         2           Platygyra sp.         PPms-1.pep         8         1           Platygyra sp.         PPms-2.pep         19*         0           Platygyra sp.         PPms-G.pep         5         0           Platygyra sp.         PPms-G.pep         13         0           Montipora sp.         RTms-1.pep         6         0           Montipora sp.         RTms-5.pep         6         0           Montipora sp.         RTms-6.pep         6         0           Montipora sp.         RTms-2.pep         6         0	sviat	Pink Pocillopora	PPd57-2ms.pep	1	QVLSPQCQYGSIFWRNSYEHENMERLQCE
Pink Pocillopora         PPd574ms.pep         2           Platygyra sp.         PPms-1.pep         8         I           Platygyra sp.         PPms-2.pep         19*         (           Platygyra sp.         PPms-E.pep         5         (           Platygyra sp.         PPms-G.pep         13         (           Montipora sp.         RTms-1.pep         6         (           Montipora sp.         RTms-5.pep         6         (           Montipora sp.         RTms-6.pep         6         (           Montipora sp.         RTms-2.pep         6         (	sviat	Pink Pocillopora	PPd57-3.pep	1	QVLSPQCQYGSFWRNSYEHENMERLQCE
Platygyra sp.         PPms-1.pep         8         B           Platygyra sp.         PPms-2.pep         19**         ()           Platygyra sp.         PPms-E.pep         5         ()           Platygyra sp.         PPms-G.pep         13         ()           Montipora sp.         RTms-1.pep         6         ()           Montipora sp.         RTms-5.pep         6         ()           Montipora sp.         RTms-6.pep         6         ()           Montipora sp.         RTms-2.pep         6         ()	sviat	Pink Pocillopora	PPd57-4ms.pep	2	QVLSPQSQYGSIYWRNSYENENMERLQCE
Platygyra sp.         PPms-2.pep         19*           Platygyra sp.         PPms-E.pep         5         C           Platygyra sp.         PPms-G.pep         13         C           Montipora sp.         RTms-1.pep         6         C           Montipora sp.         RTms-5.pep         6         C           Montipora sp.         RTms-6.pep         6         C           Montipora sp.         RTms-2.pep         6         C	sviat	Platygyra sp.	PPms-1.pep	8	RVLSPOCOYGNIFWRNSYEHENMGRLOCE
Platygyra sp.         PPms-E.pep         5         (           Platygyra sp.         PPms-G.pep         13         (           Montipora sp.         RTms-5.pep         6         (           Montipora sp.         RTms-6.pep         6         (           Montipora sp.         RTms-6.pep         6         (	sviat	Platygyra sp.	PPms-2.pep	*61	QVLSPQYQYGSIFWRNSYENENMERLRCE
Platygyra sp.         PPms-G.pep         13         C           Montipora sp.         RTms-1.pep         6         C           Montipora sp.         RTms-5.pep         6         C           Montipora sp.         RTms-2.pep         6         C	sviat	Platygyra sp.	PPms-E.pep	5	QVLSPQCQYGNIFWRNSYEHENMGRLQCE
Montipora sp.         RTms-1.pep         6         0           Montipora sp.         RTms-5.pep         1         0           Montipora sp.         RTms-2.pep         6         0	sviat	Platygyra sp.	PPms-G.pep	13	QVLSPQCQYGNIFWGNSYEHENMGRLQCE
Montipora sp.         RTms-5.pep         1         0           Montipora sp.         RTms-6.pep         6         0           Montipora sp.         RTms-2.pep         6         0	sviat	Montipora sp.	RTms-1.pep	9	QVLSPQYQYGSIYWRNSYENENMERLQCE
Montipora sp.RIms-6.pep6(C) <td>sviat</td> <td>Montipora sp.</td> <td>RTms-5.pep</td> <td>1</td> <td>QVLSPQCQYGSIFWRNSYEHENMERLQCE</td>	sviat	Montipora sp.	RTms-5.pep	1	QVLSPQCQYGSIFWRNSYEHENMERLQCE
Montipora sp. RTms-2.pep 6	svivt	Montipora sp.	RTms-6.pep	9	QVLSPQYQYGSIYWRNSYENENMERLQCE
	svsat	Montipora sp.	RTms-2.pep	9	QVLSPQYQYGSIYWRNSYBNENMERLQCE

ophore	TERLOCE	TERLORE	TERLOCE	GRLQCE	TERLORE	IERLQCE	FRLQRE	AERLQCE	<b>TESIQCE</b>	ERLOCE	ERLOCE	IERLQCE	TERLOCE	TERLOCE	TERLOCE	TERLORE	ERLORE	ERLQRE	AERLQCE	TERLOCE	MERLQRE	TERLOCE	TERLOCE	*MNE	TERLOCE	IGRLQCE	IGRLOCE	,
Amino acids within 5. A of fluorophore	QVLSPQSQYGSIYWRNSYENGNMERLQCE	QVLSPQSQYGSIYWRNSYENENMERLQRE	QVLSPQSQYGSIYWRNSYENENMERLQCE	QVLSPRCQYGNIFWRNSYEHENMGRLQCE	QVLSPQSQYGSIYWRNSYENENMERLQRE	QVLSPQCQYGSIFWRNSYEHENMERLQCE	QVLSPQSQYGSIYWRNSYENENMERLQRE	QVXSPQSQYGSXYWRNSYEHENMERLQCE	QVLSPQCQYGSIFWRNSYEHENMESIQCE	QVLSPQCQYGSIFWRNSYEHENMERLQCE	QVLSPQSQYGSTYWRNSYENENMERLQCE	QVLSPQSQYGSIYWRNSYENENMERLQCE	QVLSPQCQYGSIFWRNSYEHENMERLQCE	QVLSPQSQYGSTYWRNSYENENMERLQCE	QVLSPQSQYGSIYWRNSYENENMERLQCE	QVLSPQSQYGSIYWRNSYENENMERLQRE	QVLSPQSQYGSIYWRNSYENENMERLQRE	QVLSPQSQYGSIYWRNSYENENMERLQRE	QVLSPQSQYGSVYWRNSYVNENMERLQCE	QVLSPQCQYGSIFWRNSYEHENMERLQCE	QVLSPQSQYGSVYWRNSYENENMERLQRE	QVLSPQSQYGSTYWRNSYENGNMERLQCE	QVLSPQSQYGSIYWRNSYENENMERLQCE	QVLSPQSQYGSIYWRNSYENENM*	QVLSPQSQYGSTYWRNSYENENMERLQCE	QVLSPQCQYGNIFWRNSYEHENMGRLQCE	QVLSPQCQYGNIFWRNSYEHENMGRLQCE	
Type	15	14	2	4	14		14	20	18		2	2	1	2	2	14	14	14	7	. 1	17	15	2	2*	2	5	5	
Name	Aasv-1.pep	Aasv-3.pep	Aasv-P.pcp	Acasv-A.pep	Acasv-C.pep	Acasv-D.pep	Ce61-3sv.pep	Ce61-4sv.pep	Ce61-5sv-rep.pep	Ce61-7sv-rep.pep	GPd58-2sv.pep	LGAsv-A.pep	LGAsv-C.pep	LGAsv-D.pep	LGAsv-E.pep	Misv-A.pep	Misv-B.pep	Misv-F.pep	Pavsv-A.pep	Pavsv-B.pep	Pavsv-C.pep	PM1Asv-rep.pep	PM1Csv-rep.pep	PMsv-4.pep	PMsv-5.pep	PPsv-1.pep	PPsv-2.pep	
Species	Acropora aspera	Acropora aspera	Acropora aspera	Acanthastria sp.	Acanthastria sp.	Acanthastria sp.	Caulastrea ap.	Caulastrea ap.	Caulastrea ap.	Caulastrea ap.	Green Pocillopora	Acropora nobilis	Acropora nobilis	Acropora nobilis	Acropora nobilis	Millepora sp. (Hydrozoan)	Millepora sp. (Hydrozoan)	Millepora sp. (Hydrozoan)	Pavona decussaca	Pavona decussaca	Pavona decussaca	Porites Murrayensis	Porites Murrayensis	Porites Murrayensis	Porites Murrayensis	Platygyra sp.	Platygyra sp.	
N-terminus	sviak	sviak 1	sviak /	sviak   /	sviak	sviak	sviak	sviak	sviak	sviak	sviak	sviak	sviak															

N-terminus	Species	Name	Type	Amino acids within 5 Å of fluorophore
sviak	Platygyra sp.	PPsv-4.pep	4	QVLSPRCQYGNIFWRNSYEHENMGRLQCE
sviak	Platygyra sp.	PPsv-5.pep	2	QVLSPQSQYGSIYWRNSYENENMERLQCE
sviak	Platygyra sp.	PPsv-6.pep	10	QVLSPQSQYGSIYWRNSYENENMERLQCG
sviak	Montipora sp.	RTsv-1.pep	2	QVLSPQSQYGSIYWRNSYENENMERLQCE
sviak	Montipora sp.	RTsv-2.pep	2	QVLSPQSQYGSIYWRNSYENENMERLQCE
sviak	Montipora sp.	RTsv-3.pep	2	QVLSPQSQYGSIYWRNSYENENMERLQCE

- 151 -

TABLE 8 Percentage DNA sequence similarities generated using LALIGN

	A8	D1	D10	S1	S3	<b>T1</b>	T3
A8		97.3	98.7	97.7	99.6	97.5	99.9
D1			97.5	98.1	96.9	99.6	97.2
D10				98.2	98.2	97.6	98.5
S1_					97.9	98.2	97.5
· S3						97.0	99.4
T1							97.3
T3							

TABLE 9 Percentage amino acid sequence similarities generated using LALIGN

:	A8	D1	D10	S1	S3	<b>T</b> 1	T3
A8		95.5	98.2	97.3	99.1	96.0	100.0
D1			94.6	96.4	94.6	98.7	95.5
D10				96.4	97.3	95.1	98.2
S1					97.3	96.9	97.3
S3						95.1	99.1
T1							96.0
T3							

# $extsf{TABLE}$ 10

Mins-A=Mins-B Aams-2=Aams4

Acasv-D = PavsvB PPd57-2ms=PPd57-3

Blue: \max = 589 - 593 rm Type 1: QW.SPQCQYGSIFWRNSYEHENMERLQCE

- 152 -F F F F F F F F F F 0 0000000 **> > >>>>>>** > ≥ 3 3 3 3 3 3 3 z zzzzz TIIIIII I z z ∢ 444444 Ø Ø 0 000000 0 0 0 0 0 0 0 0 0 REPRETER IN *თ* თ თ თ თ თ თ თ თ Ø 0 0000000 KGGPL ר רורור **ග ග ග ග ග ග ග ග ග** O **७ ७७७७७७ ७** G N N F M A L K L E G O N N F M A M K L E G O N N F M A L K L E G G O N N F M A L K L E G G O N N F M A L K L E G G O N N F M A L K L E G G O N N F M A L K L E G G O N N F M A L K L E G G O N N F M A L K L E G G O N N F M A L K L E G G O N N F M A L K L E G G O N N F M A L K L E G G O N N F M A L K L E G G O N N F M A L K L E G G O N N F M A L K L E G G O N N F M A L K L E G G O N N F M A L K L E G G ο, ဟ ဟ တတတတတတ Ø < < <<<<<<<<<<<< 0 >>>>> > > > > F M A L K L F M A W K K L 0z z ZZZZZZ ¥ Ø  $\omega$   $\omega$   $\omega$   $\omega$   $\omega$ **a a a a a** α. Δ. EQTVKLTVT **⊢** > >>>>> > ¥ ススススス ΥY **-}- }α** α α α α α α α α α α **)** ပ OO 000000 ⋖ **4 4 4 4 4 4 4 個** ے > 4 > > 4 4 >>>>> × ⋖ 44444 **တ တ တ တ တ တ တ တ တ** > -0 000000 Ø 0 0 0 0 0 0 0 0 Ø ۵ ۵۵ 00000 -------w ب ပ 0 0 0 0 0 0 0 0 w u, டட **u** u u u u u σ σ agaga ga Ø Ø z w ш мп пп пп пп w ш Σ > >>>> >> > > \_ \_ -ď ٠ ۵.  $\alpha$   $\alpha$   $\alpha$   $\alpha$   $\alpha$   $\alpha$   $\alpha$   $\alpha$   $\alpha$   $\alpha$ . . . . . . . . Ø **0 000000000** 0000000 XXXXXX X ¥ × и и и и и и и и и **++++ ≻≻≻≻ }-**۲ **--**Ø O ≥ < < <<<<<<<< > > >> 0000000 XXXXXXXX |-|L ¥ ¥ 0 0000000 L L LLLLLLL L Ω G O XXXXX ZZZZZ IIIII ¥ ¥ **ス**ス Ω Ø Ø Ø 0 0 0 0 0 0 z z ZZ <u>o</u> w  $\mathbf{m}$   $\mathbf{m}$   $\mathbf{m}$   $\mathbf{m}$   $\mathbf{m}$   $\mathbf{m}$   $\mathbf{m}$   $\mathbf{m}$   $\mathbf{m}$ ш ш ппппппп п Ξ I யய ۵. z z ZZZZZ ZZ u. u. U. D V T ⊢ > >>>>> H H ш m o Ø Ø 0 0 0 0 0 0 0 0 > > m marane m σ a a aaaaaa ۵ 00000 ≻ H > . ⊗ ¥ X X x x x x x x x\_ MOM TYKVYMSGTVNGH MOM TYKVYMSGTVNGH MOM TYMSYMSGTV TQM TYKVYMSGTV TQM TYKVYMSGTV TQM TYKVYMSGTV I > ¥ 7 V N G Ø ტ 9999999 a σ æ KVYMSGTVN N A L A R A A R B B A A A B **Y ≻** ٥ H Y ∨ [ \*\*\*\* ¥ x x x x x x x x xØ 9 9 9 9 9 9 9 9 Ø YMS >>>>>>>>>>> Σ ≻ > >>>>>> Σ Σ ဖဖ m m m m m m m m m m> > <u>-</u> ٥. Ω, ••••• > Q. O. P V K M ტ Ø Ö QMTY Q M T X zz ZOZZZZ z ο. ο. **a a a a a a a a** -۵. KAKK. u. 0 0 0 000000 4 4 4 4 4 4 4 4 0 C C KTTYKAK -----> >>>>> > R R S HVKFS エイベ တတတတတတ တ 'n S တတတတတတ S шишици ц 591 T Y H V K F 591 T Y H V K F 593 T Y H V K F 589 T Y H V K F 583 M Y H V K F 583 M Y H V K F Ö Ö 0 000000 a a a aaaaaa -> = ᅶ XXXXX $\times \times$ Ŀ. և և և և և և 7 **თ თ თ თ თ თ თ თ** S ш **ш пппп**п 591.5 L 591.5 L 593.5 \_ ပ 0 00000 00 8 59.5 588 889 393 32.5 293 591.5 583 583 292 ) } acue Blue Blue Clear Blue Blue CEAR CLEAR BLUE BLUE BLUE CLEAR BLUE BLUE BLUE BLUE BLUE BLUE BLUE CLEAR BLUE BLUE Q.EAR FBLUE BLUE BLUE BLUE CLEAR Acasy-D P(61Lper Casy-D P161L per Acasy-D P161L, pet Mms-A.pep Mms-B.pep PPd57-2ms.pep Aams-2.pep Ce61-7sv.pep Mirrs-A.pep Mirrs-B.pep PPG57-2ns.pep Aarrs-2.pep LGAsv-C.pep Ce61-7sv.pep Mims-A.pep Mims-B.pep LGAsv-Cpep Ce61-7sv.pep Mirrs-Apep Mirrs-B.pep PPd57-2rts.pep LGASA-C.pep Ce61-7sv.pep RTms5.pep Acasv-0.pep Acasy-D P161 RTms5.pep Acasv-D.pep -GASA-C.pep RTms5pep Acasv-D.pep KTmst.pep Acasv0.pep Aams-2.pep ams-2.pep eq ID No: 30 (eq ID No: 34 (eq ID No: 38 (eq ID No: 122 pq ID No: 122 pq ID No: 140 pq eq ID No: 140 pq eq ID No: 88 fq ID No: 165 eq ID No: 30

ΗYL

RDGMLIGNNFMALKLEGGG

# TABLE 11

MSVIANTANTYKVYMSGTVNGHYFEWEGDGKGKPYEGEOTVKLTVTKGGPLPFAWDIMSVIAMGMTYKVYMSGTVNGHYFEVEGDGKGKPYEGEOTVKLTVTKGGPLPFAWDIMSVIAMGMTYKVYMSGTVNGHYFEVEGDGKGKPYEGEOTVKLTVTKGGPLPFAWDIMSVIAMGMTYKVYMSGTVNGHYFEVEGDGKGKPYEGEOTVKLTVTKGGPLPFAWDIMSVIAMGMTYKVYMSGTVNGHYFEVEGDGKGKPYEGEOTVKLTVTKGGPLPFAWDIMSVIAMGMTYKVYMSGTVNGHYFEVEGDGKGKPYEGEOTVKLTVTKGGPLPFAWDIMSVIAMGMTYKVYMSGTVNGHYFEVEGDGKGKPYEGEOTVKLTVTKGGPLPFAWDIMSVIATUMTYKVYMMGTVNGHYFEVEGDGKGKPYEGEOTVKLTVTKGGPLPFAWDIMSVIATUMTYKVYMMGTVNGHYFEVEGDGKGKPYEGEOTVKLTVTKGGPLPFAWDIMSVIATUMTYKVYMSGTVNGHYFEVEGDGKGKPYEGEOTVKLTVTKGGPLPFAWDIMSVIATUMTYKVYMSGTVNGHYFEVEGDGKGKPYEGEOTVKLTVTKGGPLPFAWDIMSVIATUMTYKVYMSGTVNGHYFEVEGDGKGKPYEGEOTVKLTVTKGGPLPFAWDIMSVIATUMTYKVYMSGTVNGHYFEVEGDGKGKPYEME RISv2=Alsv3 gtCP from Gusicaya.et al. 2001 FEBS Letters 607 PMTS-B = PMTS-E = PPUST-4ms LGAS-A = LGAS-N = GPUSB-23/ = PPS-6 Type 2: CALSPOSOYGSIYMRNISYBJENNER, QCE 579 280 Į purple print LGASVE pep LGASAD.pep PMCsv.pep PMT8-Apep RTsv-2 pep PMsv-6.pep PMms-Bpep Risk-Lpep

Seq ID Noz 84 Seq ID Noz 84 Seq ID Noz 82 Seq ID Noz 82 Seq ID Noz 128 Seq ID Noz 128

Sed ID No. 217

Seq 10 No.82 Seq ID Nor48

LSPCSQVGSIPFTKYPEDI--PDYVKQSFPEGYTWERIMNFEDGAVCTVSNDSSIQGNCFLSPCSQVGSIPFTKYPEDI--PDYVKQSFPEGYTWERIMSFEDGAVCTVSNDSSMQGNCFLSPCSQVGSIPFTKYPEDI--PDYVKQSFPEGYTWERIMNFWGDAVCTVSNDSSIQGNCFLSPCSQVGSIPFTKYPEDI--PDYVKQSFPEGYTWERIMNFWDGAVCTVSNDSSIQGNCFLSPCSQVGSIPFTKYPEDI--PDYVKQSFPEGYTWERIMNFEDGAVCTVSNDSSIQGNCFLSPQSQVGSIPFTKYPEDI--PDYVKQSFPEGYTWERIMNFEDGAVCTVSNDSSIQGNCFLSPCSQVGSIPFTKYPEDI--PDYVKQSFPEGYTWERIMNFEDGAVCTVSNDSSIQGNCFLSPCSQVGSIPFTKYPEDI--PDYVKQSFPEGYTWERIMNFEDGAVCTVSNDSSIQGNCF SIPFTKYPED! - . PDY V KQSFPEGYTWER! MN FEDGAVCT VS N DSSI Q GNCF SIPFTKYPEDI...PDYVKQSFPEGYTWERIMN FEDGAVCTVS NDSSIQGNCF QYGSIPFTKYPEDI...PDYVKQSFPEGYTWERIMN FEDGAVCTVS N DSSIQGNCF 91 LSPGS 91 LSPGS 91 LSPGS 91 LSPGS ST LSPCS LSPCS 6 5 579.5 579 88 deze pupe pupe purple purple å ž

115 I Y H V K F S G L N F P P N G P V M • Q K K T Q G W E P N T E R L F A • • • R D G M L I G N N F M A L K L E G G G RTS+1,pep

РМте-Арер РМте-В.рер

PMCsv.pep

RTsv-2pep Flyby-6.pep

-GASA-D.pep

GASAEpen

115 I YHVK F S G L N F P P N G P V M - Q K K T Q G W E P N T E R L F A - .

ğ

LGASA-Epep CGASA-O,Dep PMCskpep

КТви-2рер

579.141 YHVKFSGLNFPPNGPVM-QKKTQGWEPNTERLFA-・RDGMLIGNNFMALKLEGGG 579.5 116 IYHVKFSGLNFPPCGPVM-QKKTQGWEPNTERLFA-・RDGMLIGNNFMALKLEGGG 115 IYHVKFSGLNFPPNGPVM-QKKTQGWEPNTERLFA-・RDGMLIGNNFMALKLEGGG 579 115 IYHVKFSGLNFPPNGPVM-QKKTQGWEPNTERLFA-・RDGMLIGNNFMALKLEGGG 116 IYHVKFSGLNFPPNGPVM-QKKTQGWEPNTERLFA-・RDGMLIGNNFMALKLEGGG 579.5 116 IYHVKFSGLNFPPNGPVM-QKKTQGWEPNTERLFA-・RDGMLIGNNFMALKLEGGG 580 115 I YNVK F SG LNF P PNGPVM - QKKTQGWEPNTER L FA - - RDG ML I purple dear dear purple purple 뛆 Berga

PMms-Apep PMms-Bpep

- Avave pep

H Y L H Y L H Y L H Y L

LEGGG

GNNFMALK

17 CEFKS T Y K AKK - PVKMPGYHYVDR K LDVTNHNKDYTS. V E QCE I S I A RKP TH CEFKS TYKAKK - PVKMPGYHYVDRK LOVTNHNKDYTS - VE QCE I SI ARKP perpe ğ Семеновр GASVEDED PMC<sub>6</sub>v.pep RTSv-1 pap RTsw2pep

579 170 CEFKS T Y K AKK - PVKMPGYHYVDRKLDVTNHNKDYTS - V E GCE I S I ARKP V V A 579 - S 171 CEFKS T Y K AKK - PVKMPGYHYVDRKLDVTNHNKDYTS - V E GCE I S I ARKP V V A 172 CEFKS T Y KAKK - PVKMPGYHYVDRKLDVTNHNKDYTS - V E GCE I S I ARKP V V A 579 171 CEFKS T Y KAKK - PVKMPGYHYVDRKLDVTNHNKDYTS - V E GCE I S I ARKP V V A 579 171 CEFKS T Y KAKK - PVKMPGYHYVDRKLDVTNHNKDYTS - V E GCE I S I ARKP V V A 171 CEFKS T Y KAKK - PVKMPGYHYVDRKLDVTNHNKDYTS - V E GCE I S I ARKP V V A 171 CEFKS T Y KAKK - PVKMPGYHYVDRKLDVTNHNKDYTS - V E GCE I S I ARKP V V A 579.5 171 CEFKS TYKAKK • PVKMPGYHYVDRK LDVTNHNKDYTS • VE QCE I SI ARKP V VA 170 CEPKS T Y KAKK . PVKMPGYHYVDRKLDVTNHN FOTTS . VE QCE 1 S I ARKP V VA 288 purple purple desir purple purple

> PMTE-Appp PAMPE E pep

4	
<del></del>	
Ħ	
긆	
AB	
Þ	
Η	

Auma a 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	F KOTYKARK • PVKRPGYHY VORKLOVINHNK DYIM• VM	VORKLOV1NHNKOV18	FKSTYKARK-PVKMPGYKWWORKLOVINHNKDYTS.VEQR FKSTYKARK-PVKMPGYHVVORKIDVINHNKDYTS.VEGR	EFKSTYKARK.PVKMPGYHYVDRKLDV1NHNKDY1S-VEQRE	YHVKFSGLNFPPNGPVM、QKKTQGWEPNTERLAA - RDGML! YHVKFSGLNFPPNGPVM、QKKTQGWEPNTERLAA - RDGML! YHVKFSGLNFPPNGPVM、QKKTQGWEPNTERLAA - RDGML! YHVKFSGLNFPPNGPVM、QKKTQGWEPNTERLAA - RDGML! YHVKFSGLNFPPNGPVM、QKKTQGWEPNTERLFA - RDGML! YHVKFSGLNFPPNGPVM - QKKTQGWEPNTERLFA - RDGML!	SPOSON GSIPTKYPEDIPOYVKQSFPEGYTWERIMNFEDSPOSON GSIPTTKYPEDIPOYVKQSFPEGYTWERIMNFEDSPOSON GSIPTTKYPEDIPOYVKQSFPEGYTWERIMNFEDSPOSON GSIPTTKYPEDIPOYVKQSFPEGYTWERIMNFEDSPOSON GSIPTTKYPEDIPOYVKQSFPEGYTWERIMNFEDSPOSON GSIPTTKYPEDIPOYVKQSFPEGYTWERIMNFED	V E G D G K G K P Y E G E Q T V E G D G K G K P Y E G E Q T V E G D G K G K P Y E G E Q T V E G E Q T	SV ! A KOMTYKVYMSGTVNGHYFEVEGDGKGKPYEGEGT SV ! A KOMTYKVYMSGTVNGHYFEVEGDGKGKPYEGETT	
Amma and and and and and and and and and an	O	ပပ	ပပ	ပ					
ep purple  pur	579 1			-	579,5		579.5	579.5	n F
	proble	Program	purple pink	ejama	purple purple purple	purpio purpio purpio	purple purple purple	purple	
Seq ID No:22 Assv-3.p Seq ID No:23 Assv-3.p Seq ID No:22 Assv-3.p Assv-3.p Misv-Apc Misv-Apc Assv-3.p Assv-3.p Assv-3.p Assv-3.p Assv-3.p Assv-3.p Assv-3.p Assv-3.p Assv-3.p	Msv-8.pep	Ce61-3sv.pep pink Misv-A pep purple	AcaevCpep Ce61-3sv.pep		Assv3.pep purple Acasv-C.pep purple Ce61-3sv,pep pirk Misv-Apep purple Misv-Apep purple	Assivation purpo Acasivation purpo Celivation purpo Man-Apo purpo Man-Apo purpo	Seq ID No.23 Acts W.C.pop 18 Seq ID No.22 Ce81-3sv.pep 18 Seq ID No.50 MiswApep 18 Seq ID No.52 MiswApep 18	Seq ID No:22 Assv-3.pep purple Seq ID No:28 Acastv-C.pop purple	

3
Ţ
뉙
크
Ŋ
7
ſ
1

Type 6: QVLSPQYQYGSIYWRNSYENENMERLQCE

MSVIATQMTYKVYMSGTVNGHYFEVEGDGKGKPYEGEQTVRLTVTKGGPLPFAWDIMSVIATQMTYKVYMSGTVNGHYFEVEGDGKGKPYEGEQTVKLTVTKGGPLPFAWDIMSVIATQMTYKVYMSGTVNGHYFEVEGDGKGKPYEGEQTVKLTVTKGGPLPFAWDIMSVIATQMTYKVYMSGTVNGHYFEVEGDGKGKPYEGEQTVKLTVTKGGPLPFAWDIMSVIATQMTYKVYMSGTVNGHYFEVEGDGKGKPYEGEQTVKLTVTKGGPLPFAWDIMSVIATQMTYKVYMSGTVNGHYFEVEGDGKGKPYEGEQTVKLTVTKGGPLPFAWDI	LSPQYQYGSIPFTKYPEDIPDYVKQ FPEGYTWERIMNFEDGAVCTVSNDSSIQGNCFLSPQYQYGSIPFTKYPEDIPDYVKQ FPEGYTWERIMNFEDGAVCTVSNDSSIQGNCFLSPQYQYGSIPFTKYPEDIPDYVKQ FPEGYTWERIMNFEDGAVCTVSNDSSIQGNCFLSPQYQYGSIPFTKYPEDIPDYVKQ FPEGYTWERIMNFEDGAVCTVSNDSSIQGNCFLSPQYQYGSIPFTKYPEDIPDYVKQ FPEGYTWERIMNFEDGAVCTVSNDSSIQGNCFLSPQYQYGSIPFTKYPEDIPDYVKQ FPEGYTWERIMNFEDGAVCTVSNDSSIQGNCF	IYHVKFSGLNFPPNGPVM-QKKTQGWEPNTERLMA-、RDGML!GNNFMALKLEGGG HYLIYHVKFSGLNFPPNGPVM-QKKTQGWEPNTERLFA-・RDGML!GNNFMALKLEGGG HYLIYHVKFSGLNFPPNGPVM-QKKTQGWEPNTERLFA-・RDGML!GNNFMALKLEGGG HYLIYHVKFSGLNFPPNGPVM-QKKTQGWEPNTERLFA-・RDGML!GNNFMALKLEGGG HYLIYHVKFSGLNFPPNGPVM-QKKTQGWEPNTERLFA-・RDGML!GNNFMALKLEGGG HYLIYHVKFSGLNFPPNGPVM-QKKTQGWEPNTERLFA-・RDGML!GNNFMALKLEGGG HYLI	: EFKSTYKAKK - PVKMPGYHYVDRKLDVTNHNKDYTS - VEQCE I SIARKPVVA
	57 57 57 57	115 116 115 113	- 171
583.5 584 584.5	583.5 584 585.5	583.5 584 585.5	583.5 171
purple purple dear purplo clear	purpla purpla clear purpla clear	purple purple clear purple dear	purple
Seq ID Nor116 LGArrs-S.pep purple Seq ID Nor162 RTms-1.pep purple Seq ID Nor158 Pavrrs-3.pep dear Seq ID Nor168 RTms-6.pep purplo Seq ID Nor164 RTms-2.pep dear	LGArrs-Spep purple RTrrs-1.pep purple Pavrrs-3.pep clear RTrrs-6.pep purple RTrrs-2.pep dear	LGAms-Spep purple RTms-1.pep purple Pavms-3.pep clear RTms-6.pep purple RTms-2.pep dear	LGAnts-5.pep purple
Seq ID Nor116 LGAnra-G.pep Seq ID Nor162 RTrns-1.pep Seq ID Nor158 Pavrns-3.pep Seq ID Nor164 RTrns-G.pep Seq ID Nor164 RTrns-C.pep			

T. TABLE 14

Sequence types with no 520-600 nm absorbance producing non-coloured bactaria (clear)

name	٤	Amino 88	acdd (a	Is with	nin Sr	lm of	chromo	Amino acids within 5 ran of chromophore with M of MSVIA≖ 1: 39 40 57 68 50 60 64 65 65 65 65	with Mo	of MSM	Ω=1:		5														prodominant coral pigment	
PPsv-1.pep	2		1		3 02	1	ı	▝	3 >	ğ	8 - 3 -	اۋ	3	5 0	2	٠ŧ	2   2   4   4	וב	٦,	Τ,	盟:	][	۱.	28 28 28		五年 210		1
P9sv-2.pep	U.	c	· >	-	0	۵.		\$ (	;	9 (	- ·	ĻI	> }	۲ (		ֶ ה	≻ ח	.,	IJ	z	Σ	IJ	r	a a	~	m	green fluorescent	
DDW. 1-VOD	u	3 (	> >	J .	2 0	٠.	,,	7 (		٠	z :	1	3	Ľ	z	S	w ≻	I	Ш	z	Σ	Ø	~	ں	~	w	green fluorescent	
Course C non	טכ	3 (	> ;	. د	0	ı, a	J (	ġ (	<b>&gt;</b>	IJ.	z	u.	≥	œ	z	Ś	w ≻		w	z	≆	O	æ	_,	~	W	green fluorescent	
dad H-sug	o u	3 (	> >	_ ر	n o	ı.	J (	ġ (	>	ڻ ا	z:	u.	≥	œ	z	'n	ш >-		w	z	Σ	O	~	ں۔	0	W	green fluorescent	
Phys-E.pen	y c	3 (	> >	J .	<i>o</i> 0	r c	30	<i>3</i> (	<b>,</b> :	ۍ ن	- : z :	<b>L</b>	≥ :	œ	z	s	ш Ж	I	ш	z	Σ	Ø	æ		ن ~		green fluorescent	
Coma-P. nen	, u	3 (	> >	_ د	, (		, c	3 (	> } }	9	- ·		≥ .	œ;		s s	.ω ~		w	z	≥	U	œ	g _		m m	green fluorescent	
Dems-G-pen	, ţ	3 (	> >	J _	0 0	L C	J (	3 (	<b>&gt;</b> ;	9 (	- z:	<b>L</b> . 1	≥ "	œ	z	S	γ Ш		w	z	Σ	တ	æ	a			green fluorescent	
DPms-1.nen	2 ∝	מינ	> >	د	0 0	L 0	י טיפ	•	<b>,</b>	ۍ د	 Z :	1	≥		z	S	ш >-	<b>I</b>	w	z	Σ	တ	~	a	o ~	m	green fluorescent	
Araguel nor	•		> >	J.	0 0	r (	יינ ער		<b>&gt;</b> :	ŋ	z	u.	₹	œ	z	S	m	Ξ	ш	z	×	O	æ	σ,	0	ш	green fluorescent	
Bear A non	٠,	3 (	> >	٠.	0	kiin L (	) (四	3 (	<b>&gt;</b>	<u>ق</u>	z	u.	≥	œ	z	≻ ທ	~ ш	Œ	w	z	≥		æ	a	0		green fluorescent	
Total Laboratory	, 4	3 (	> :	. د	י מ	iii. L	関	3	>	ن ن	- z	<b>և</b>	≥	œ	z	r vo	ъ Б	Ξ	w	z	Σ	ဟ	2	σ,	٥	ш	green fluorescent	
dad - Taken	2 !	3 (	> ;	. د	o	a.		o Hill	➤	<sub>ග</sub>			≥	œ	z	≻ ທ	w ≻		関節	z	###  ≥		2	σ.		ш	theosard floor	
recovery.	₽	σ	>	ے	v	۰	O S	CO IIII	<b>,</b>	Ü	-	S.	3	œ	z	<i>&gt;</i>	ц.		肥泥	z	IHI ≤	1	0			u		
Pavsv-C.per	11	σ	>	د	S	a.		O	>	O O			3	α	: z	· >			ш	2	###i		- د د د	,	1	9 L	bupie	
Pavev-A.Dec	7	C	>	_	v	۵		C HILL	>	(	THE REAL PROPERTY.			: 6	: :	- :	)      -		J.	Z	14.C E	J.	٠	,		U M	green monescera	
•		,		3	,		III	3		9			\$	r	z	s	>   		W	z	Σ	IJ	~	G	_ C	Œ	measuril needs	

Sequence types with some 520-600 nm absorbance producing plinky-purple bacteria

	рестипати соға рідмент деел білуғасен
	<b>事</b> m m
	800
	800
	3 4 4
	1
	\$ ≥ ≥ ≥
	2 2 2
	4.10
	名 m m
	<del>2</del> > >
	S S
	₽ZZ
	2000
	2 × 2 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
٠.	8 % -
÷ n ≤	20 00
₹	ୟ ଦ.ଦ
ž	8 > >
₹.	800
	2 図版
5	3 a a
Ē	300
	3 o o
THE SCIENCE WILL	
2	송 > >
į	2 a a
	5 8 <del>2</del>
	rama Ce61-4ev PPU57-1ms

TABLE 15 Conserved amino acid differences between blue and purple colored proteins

Position	Amino acid in blue proteins (n = 2)	Amino acid in purple proteins (n = 4)	Amino acid in blue- purple protein (n = 1)
41	Arg (charged, polar)	Lys (charged, polar)*	Arg
43	Ala (nonpolar)	Thr (uncharged, polar)	Ala
61	Cys (uncharged, polar)	Ser (uncharged, polar)	Cys
87	Phe (nonpolar)	Tyr (uncharged, polar)	Phe
142	His (charged, polar)	Asn (uncharged, polar)	Asn
143	Ser (uncharged, polar)	Thr (uncharged, polar)	Thr
175	Thr (uncharged, polar)	Ser (uncharged, polar)	Ser
198	Ile (nonpolar)	Thr (uncharged, polar)	Thr

<sup>\*</sup> Amino acid position 41 of the purple protein encoded by D10 (SEQ ID NO:192) is Arg.

**TABLE 16:** Amount of colored protein (expressed as a percentage of total soluble protein) produced in cultures of *E. coli* and *S. cerevisiae*.

Species	Plasmid	CP clone	Colour	RHSCC	%CP
E.coli	pCGP2921	T1	blue	102A	50%
S. cerevisiae	pCGP3269	A8	purple	82B	8%
S. cerevisiae	pCGP3270	T1	blue	101C	6%

RHSCC = Royal Horticultural Society Colour Chart (Kew, UK)

TABLE 17 Summary of recombinant protein accumulation levels in plants after nuclear DNA transformation.

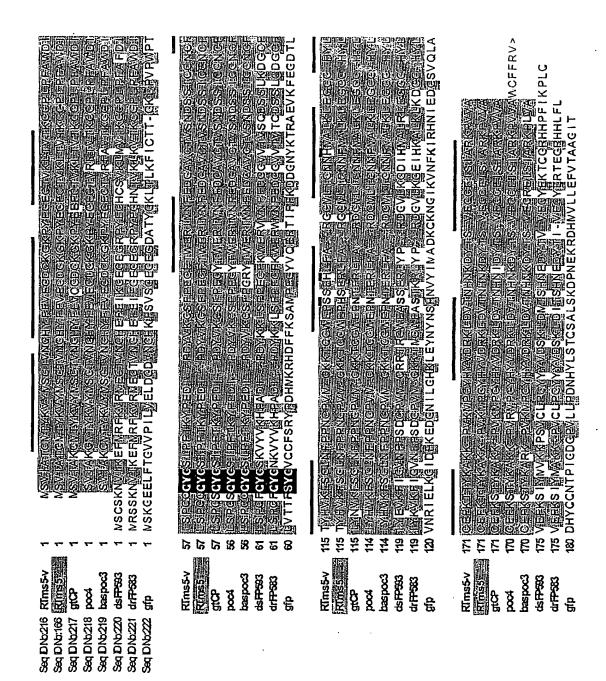
Protein	Plant	Targeting	Accumulation	Reference
Endoglucanase	Tobacco	Chloroplast (Tomato- Rubisco small subunit protein)	Up to 1.35 % TSP in leaves.	Transgenic Res 9: 43-54, 2000
PEPC	Rice	cytosol	Up to 12 % TSP in leaves.	Nat Biotechnol. 17: 22-23, 1999
Cystatin	Rice	cytosol	Up to 2 % TSP.	Plant Mol Biol 30: 149-157, 1996
Antibody	Arabid.	ER-retained (DIKDEL), ER- secreted	Up to 6 % TSP	Eur J Biochem
Spider Silk	Tobacco Potato	ER-retained	2 % + TSP in leaves and potato tubers	Nat Biotechnol 19: 573-577. 2000
Cry9Aa	Tobacco Potato Cauliflower Turnip	cytosol	0.3 % TSP in Tobacco leaves. Expression in other plants 0.1-0.03 %.	Plant Sci 160: 341- 353, 2001
Xylanase GFP Alkaline phosphatase	Tobacco	ER-excreted in guttation fluid	3 % TSP (alk. phos.)	Plant Physiol 124: 927-934, 2000
GUS- Peptide vaccine	?	cytosol	Up to 3 %TSP	FEBS Lett 488: 13- 17, 2001
Bt, NPTII	Tobacco	cytosol	Bt: 0.02% TSP NPTII: 0-07-0.27 % TSP	Nature 328: 33-37, 1987
AlMV-CP	Tobacco Tomato	cytosol	0.1 - 0.4 % TSP Tobacco 0.1 - 0.8 % TSP Tomato	EMBO 6: 1181 -1 1188, 1987
sGFP	Rice	Chloroplast (rbsS-Tp)	10 % of TSP. Much higher than cytoplasmic control (0.5 % TSP)	Plant & Animal Genome VII Conference 1999 abstract

TSP = total soluble protein

TABLE 18 Summary of recombinant protein accumulation levels in plants after Plastid DNA transformation.

Protein	Plant	Targeting	Accumulation	Reference
GFP	tobacco	Chloroplast expression	5 % TSP in leaves	Plant Journal 27: 257 - 265
Bt (cry2Aa2)	tobacco	Chloroplast expression	2-3 % TSPin leaves	Proc Natl Acad Sci 1840-1845, 1999
Bt (cry2Aa2)	?	Chloroplast expression	45.3 % TSP	Nat Biotechnol 19: 71-74, 2001
Somatotropin	Tobacco	Chloroplast expression	7 % + "more than 300- fold higher than a similar gene expressed using a nuclear transgenic approach"	Nat Biotechnol 18: 333-338, 2000
Bt (crylAc)	Tobacco	Chloroplast expression	3-5 % TSP in leaves	Biotechnology 13: 362-365, 1995
EPSPS	Tobacco	Chloroplast expression	10 % + TSP in leaves	Plant J 25: 261- 270, 2001

TSP = total soluble protein



- 161 -

 TABLE 20
 Summary of Northern analysis of Arabidopsis transgenic plants

Construct number	CP Cassette	Selectable marker	T1	SuRB
pCGP2772	35S: T1: 35S	35S: SuRB	~0.9 kb	~2.2 kb
pCGP2765	35S: A8: 35S	35S: SuRB	~0.9 kb	~2.2 kb
pCGP3259	35S: ER:T1: 35S	35S: SuRB	~1.0 kb	~2.2 kb
pCGP2785	35S: SSU:T1: 35S	35S: SuRB	~1.1 kb	~2.2 kb
pCGP3258	35S: T1:GFP: 35S	35S: SuRB	~1.6 kb	~2.2 kb
pCGP3261	35S: ER:T1:GFP: 35S	35S: SuRB	~1.7 kb	~2.2 kb
pCGP960	35S: GUS	35S: SuRB	none	~2.2 kb
pBINmgfp4	35S: mGFP4: nos	35S : <i>npt</i> ∏	none	none
non transgenic	NA	NA	none	none

CP cassette = Colored protein cassette contained in construct;

SM cassette = the selectable marker gene contained in construct;

NA = not applicable;

none = no transcripts detected

- 162 -

TABLE 21 Estimations of T1 protein in leaf samples from 2 transgenic Arabidopsis events (expressed as a percentage of total protein)

Construct	Cassette	Acc#	Sample	% <b>T</b> 1
pCGP3259	35S:ERT:T1:35S	1.5	leaf	0.005%
pCGP2772	35S:T1:35S	1.2b	leaf	0.005%

Construct = Binary vector used in transformation;

Cassette refers to the chimaeric T1 transgene contained in the T-DNA;

Acc# refers to the accession number of the transgenic plant.

TABLE 22: Estimations of T1 protein in petal and/or leaf samples from 2 transgenic P.

hybrida events (expressed as a percentage of total protein)

Construct.	Cassette	Acc#	Sample	%T1:
pCGP3259	35S:ERT:T1:35S	24444	leaf	0.009%
pCGP2765	35S:A8:35S	24534	petal	0.02%
pCGP2765	35S:A8:35S	24534	leaf	0.006%

- 163 -

## TABLE 23 Complete amino acid sequence of PdGFP-T3.pep

SYL NGIAEEMKTDL Given N-terminal polymorphy,

## Continuing...

### PdGFP-T3.pep

- 1 MEGIVDGHKF VITGHGNGNP FEGKQTMNLC VVEGGPLPFS EDILSAAFDY
- 51 GNRVFTEYPQ GMVDFFKNSC PAGYTWHRSL LFEDGAVCTT SADITVSVEE
- 101 NCFYHNSKFH GVNFPADGPV MKKMTTNWEP SCEKIIPVPR QGILKGDIAM
- 151 YLLLKDGGRY RCOFDTIYKA KSDPKEMPEW HFIQHKLTRE DRSDAKNQKW
- 201 QLVEHAVASR SALPG\*

751 ΑΑΑΑΑΑΑΑΑ ΑΑΑΑΑΑΑ

1 ATGGAAGGA TTGTCGATGG GCATAAATTT GTGATCACGG GCCACGGCAA
51 TGGAAATCCT TTCGAAGGA AACAGACTAT GAATCTGTGT GTGGTTGAAG
101 GGGGACCCCT GCCATTCTCC GAAGACATTT TGTCTGCTGC GTTTGACTAC
151 GGAAACAGGG TCTTCACTGA ATATCCTCAA GGCATGGTTG ACTTTTTCAA
201 GAATTCATGT CCAGCTGGAT ACACATGGCA CAGGTCTTTA CTCTTTGAAG
251 ATGGAGCAGT TTGCACAACT AGTGCAGATA TAACAGTGAG TGTTGAGGAG
301 AACTGCTTTT ATCACAATTC CAAGTTTCAT GGAGTGAACT TTCCTGCTGA
351 TGGACCTGTG ATGAAAAAGA TGACAACTAA TTGGGAGCCA TCCTGCGAGA
401 AAATCATACC AGTACCTAGA CAGGGGATAT TGAAAGGGGA TATTGCCATG
451 TACCTTCTTC TGAAGGATGG TGGGCGTTAT CGGTGCCAGT TCGACACAAT
501 TTACAAAGCA AAGTCTGACC CGAAAGAGAT GCCGGAGTGG CACTTCATCC
551 AACATAAGCT CACCCGGGAA GACCGCAGCG ATGCTAAGAA CCAGAAATGG
601 CAACTGGTAG AACATGCTGT TGCTTCCCGA TCCGCATTGC CCGGATAAGA
651 ACATGATATA GTTCAAACAT GTTGTTACAT GCGCCATGCT ATTACTNTGA
701 TGACAATGTA GTTCGAGCCA GGCCAGTAGC AATAAAGCAC ATTTCAANCA

- 164 -

## **CLAIMS**

- 1. An isolated nucleic acid molecule comprising a nucleotide sequence encoding a color-facilitating molecule (CFM) which, in a cell, alone or together with one or more other molecules imparts an altered visual characteristic to said cell when visualized by a human eye in the absence of excitation by extraneous non-white light or particle emission.
- 2. The isolated nucleic acid molecule of Claim 1 wherein the CFM is derived from Anemonia majano, Anemonia sulcata, Clavularia sp, Zoanthus sp, Discosoma sp (e.g. Discosoma striata), Aequorea sp (e.g. Aequorea victoria), Anthozoa sp, Cassiopea sp, (e.g. Cassiopea xamachana), Millepora sp, Acropora sp (e.g. Acropora aspera and Acropora nobilis), Montipora sp, Porites murrayensis, Pocillopora damicormis, Pavona descussaca, Acanthastrea sp, Platygyra sp or Caulastrea sp.
- 3. The isolated nucleic acid molecule of Claim 1 or Claim 2 wherein the CFM comprises an amino acid sequence in its N-terminal end selected from SVIAK (SEQ ID NO:5), (M)SVIAT (SEQ ID NO:6), SGIAT (SEQ ID NO:7), SVIVT (SEQ ID NO:8) or SVSAT (SEQ ID NO:9).
- The isolated nucleic acid molecule of Claim 3 wherein the CFM comprises an amino acid sequence selected from the list comprising SVIAT QMTY KVYM SGT (SEQ ID NO:10), SVIAT QMTY KVYM PGT (SEQ ID NO:11), SVIAT QVTY KVYM SGT (SEQ ID NO:12), SGIAT QMTY KVYM SGT (SEQ ID NO:13), SVIVT QMTY KVYM SGT (SEQ ID NO:14), SVSAT QMTY KVYM SGT (SEQ ID NO:15), SVIAK QMTY KVNM SGT (SEQ ID NO:16), SVIAK QMTY KVYM SDT (SEQ ID NO:17) and SVIAK QMTY X<sub>1</sub>X<sub>2</sub>YX<sub>3</sub> SGT (SEQ ID NO:18) wherein X<sub>1</sub>, X<sub>2</sub> and X<sub>3</sub> may be any amino acid provided that X<sub>1</sub> is not K; X<sub>2</sub> is not V; X<sub>3</sub> is not M.
- 5. The isolated nucleic acid molecule of Claim 3 or Claim 4 wherein the CFM comprises an amino acid sequence selected from the list comprising SEQ ID NOs:20, 22,

WO 02/070703

- 165 -

PCT/GB02/00928

24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178, 180, 190, 192, 194, 196, 198, 200 and 202 provided that, where the said amino acid sequence comprises the sequence SVIAK QMTY  $X_1X_2YX_3$  SGT,  $X_1$  is not lysine,  $X_2$  is not valine, and  $X_3$  is not methionine or an amino acid sequence having at least 60% similarity to any one or more of the above referenced sequences.

- 6. The isolated nucleic acid molecule of Claim 5 comprising a nucleotide sequence encoding a color-facilitating molecule (CFM), wherein the nucleotide sequence is selected from the list comprising SEQ ID NOs:19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, 131, 133, 135, 137, 139, 141, 143, 145, 147, 149, 151, 153, 155, 157, 159, 161, 163, 165, 167, 169, 171, 173, 175, 177, 179, 189, 191, 193, 195, 197, 199 and 201 or a nucleotide sequence having at least 60% similarity to one or more of the above referenced sequences or a nucleotide sequence capable of hybridizing to one of the above referenced sequences or a complementary form thereof under low stringency conditions.
- 7. The isolated nucleic acid molecule of any one of Claims 1 to 6 wherein the cell is a prokaryotic cell.
- 8. The isolated nucleic acid molecule of any one of Claims 1 to 6 wherein the cell is a eukaryotic cell.
- 9. The isolated nucleic acid molecule of Claim 8 wherein the eukaryotic cell is a mammalian animal cell.
- 10. The isolated nucleic acid molecule of Claim 8 wherein the eukaryotic cell is a non-mammalian animal cell.

- 166 -

- 11. The isolated nucleic acid molecule of Claim 10 wherein the eukaryotic cell is a plant cell.
- 12. The isolated nucleic acid molecule of Claim 11 wherein the plant cell is part of a plant callus or a whole plant.
- 13. The isolated nucleic acid molecule of Claim 12 wherein the whole plant is an ornamental or flowering plant or a part thereof.
- 14. The isolated nucleic acid molecule of Claim 13 wherein the plant part is a flower, root, leaf, stem, seed, fruit or fiber.
- 15. The isolated nucleic acid molecule of Claim 13 wherein the plant is selected from a rose, carnation, lisianthus, petunia, lily, tulip, pansy, gerbera or chrysanthemum.
- 16. The isolated nucleic acid molecule of of any one of Claims 1 to 15 wherein the CFM is a GFP or derivative or homolog thereof.
- 17. The isolated nucleic acid molecule of Claim 16 wherein the homolog of GFP is a non-fluorescent GFP.
- 18. An isolated color-facilitating molecule (CFM) comprising a polypeptide which, in a cell, alone or together with one or more other molecules imparts an altered visual characteristic to said cell when visualized by a human eye in the absence of excitation by extraneous non-white light or particle emission.
- 19. The isolated CFM of Claim 18 wherein the CFM is derived from Anemonia majano, Anemonia sulcata, Clavularia sp, Zoanthus sp, Discosoma sp (e.g. Discosoma striata), Aequorea sp (e.g. Aequorea victoria), Anthozoa sp, Cassiopea sp, (e.g. Cassiopea xamachana), Millepora sp, Acropora sp (e.g. Acropora aspera and Acropora nobilis),

- 167 -

Montipora sp, Porites murrayensis, Pocillopora damicormis, Pavona descussaca, Acanthastrea sp, Platygyra sp or Caulastrea sp.

- 20. The isolated CFM of Claim 19 wherein the CFM comprises an amino acid sequence in its N-terminal end selected from SVIAK (SEQ ID NO:5), (M)SVIAT (SEQ ID NO:6), SGIAT (SEQ ID NO:7), SVIVT (SEQ ID NO:8) or SVSAT (SEQ ID NO:9).
- The isolated CFM of Claim 20 wherein the CFM comprises an amino acid sequence selected from the list comprising SVIAT QMTY KVYM SGT (SEQ ID NO:10), SVIAT QMTY KVYM PGT (SEQ ID NO:11), SVIAT QVTY KVYM SGT (SEQ ID NO:12), SGIAT QMTY KVYM SGT (SEQ ID NO:13), SVIVT QMTY KVYM SGT (SEQ ID NO:14), SVSAT QMTY KVYM SGT (SEQ ID NO:15), SVIAK QMTY KVNM SGT (SEQ ID NO:16), SVIAK QMTY KVNM SGT (SEQ ID NO:16), SVIAK QMTY KVYM SDT (SEQ ID NO:17) and SVIAK QMTY X<sub>1</sub>X<sub>2</sub>YX<sub>3</sub> SGT (SEQ ID NO:18) wherein X<sub>1</sub>, X<sub>2</sub> and X<sub>3</sub> may be any amino acid provided that X<sub>1</sub> is not K; X<sub>2</sub> is not V; X<sub>3</sub> is not M.
- 22. The isolated CFM of Claim 21 wherein the CFM comprises a polypeptide having an amino acid sequence selected from the list comprising SEQ ID NOs:20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178, 180, 190, 192, 194, 196, 198, 200 and 202 provided that, where the said amino acid sequence comprises the sequence SVIAK QMTY X<sub>1</sub>X<sub>2</sub>YX<sub>3</sub> SGT, X<sub>1</sub> is not lysine, X<sub>2</sub> is not valine, and X<sub>3</sub> is not methionine or an amino acid sequence having at least 60% similarity to any one or more of the above referenced sequences.
- 23. The isolated CFM of Claim 18 wherein the cell is a prokaryotic cell.
- 24. The isolated CFM of Claim 18 wherein the cell is a eukaryotic cell.

- 168 -

- 25. The isolated CFM of Claim 24 wherein the eukaryotic cell is a mammalian animal cell.
- 26. The isolated CFM of Claim 24 wherein the eukaryotic cell is a non-mammalian animal cell.
- 27. The isolated CFM of Claim 26 wherein the non-mammalian animal cell is a plant cell.
- 28. The isolated CFM of Claim 27 wherein the plant cell is part of a plant callus or a whole plant.
- 29. The isolated CFM of Claim 28 wherein the whole plant is an ornamental or flowering plant or a part thereof.
- 30. The isolated CFM of Claim 29 wherein the plant part is a flower, root, leaf, stem, seed, fruit or fiber.
- 31. The isolated CFM of Claim 29 wherein the plant is selected from a rose, carnation, lisianthus, petunia, lily, tulip, pansy, gerbera or chrysanthemum.
- 32. The isolated CFM of any one of Claims 18 to 31 wherein the CFM is a GFP or derivative or homolog thereof.
- 33. The isolated CFM of Claim 32 wherein the homolog of GFP is a non-fluorescent GFP.
- 34. An isolated cell wherein said cell or a parent cell is genetically modified to enable the production of a color-facilitating molecule (CFM) which alone or together with one or more other molecules imparts an altered visual characteristic to said cell when visualized by a human eye in the absence of excitation by extraneous non-white light or

- 169 -

particle emission.

35. The isolated cell of Claim 34 wherein the CFM is derived from Anemonia majano, Anemonia sulcata, Clavularia sp, Zoanthus sp, Discosoma sp (e.g. Discosoma striata), Aequorea sp (e.g. Aequorea victoria), Anthozoa sp, Cassiopea sp, (e.g. Cassiopea xamachana), Millepora sp, Acropora sp (e.g. Acropora aspera and Acropora nobilis), Montipora sp, Porites murrayensis, Pocillopora damicormis, Pavona descussaca, Acanthastrea sp, Platygyra sp or Caulastrea sp.

- 36. The isolated cell of Claim 35 wherein the CFM comprises an amino acid sequence in its N-terminal end selected from SVIAK (SEQ ID NO:5), (M)SVIAT (SEQ ID NO:6), SGIAT (SEQ ID NO:7), SVIVT (SEQ ID NO:8) or SVSAT (SEQ ID NO:9).
- The isolated cell of Claim 36 wherein the CFM comprises a polypeptide having an amino acid sequence selected from the list comprising SVIAT QMTY KVYM SGT (SEQ ID NO:10), SVIAT QMTY KVYM PGT (SEQ ID NO:11), SVIAT QVTY KVYM SGT (SEQ ID NO:12), SGIAT QMTY KVYM SGT (SEQ ID NO:13), SVIVT QMTY KVYM SGT (SEQ ID NO:14), SVSAT QMTY KVYM SGT (SEQ ID NO:15), SVIAK QMTY KVNM SGT (SEQ ID NO:16), SVIAK QMTY KVYM SDT (SEQ ID NO:16), SVIAK QMTY KVYM SDT (SEQ ID NO:17) and SVIAK QMTY X<sub>1</sub>X<sub>2</sub>YX<sub>3</sub> SGT (SEQ ID NO:18) wherein X<sub>1</sub>, X<sub>2</sub> and X<sub>3</sub> may be any amino acid provided that X<sub>1</sub> is not K; X<sub>2</sub> is not V; X<sub>3</sub> is not M.
- 38. The isolated cell of Claim 36 or 37 wherein the CFM comprises a polypeptide having an amino acid sequence selected from the list comprising SEQ ID NOs:20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178, 180, 190, 192, 194, 196, 198, 200 and 202 provided that, where the said amino acid sequence comprises the sequence SVIAK QMTY X<sub>1</sub>X<sub>2</sub>YX<sub>3</sub> SGT, X<sub>1</sub> is not lysine, X<sub>2</sub> is not valine, and X<sub>3</sub> is not methionine or an amino acid sequence having at least 60% similarity to any

- 170 -

one or more of the above referenced sequences.

- 39. The isolated cell of Claim 34 wherein the cell is a prokaryotic cell.
- 40. The isolated cell of Claim 34 wherein the cell is a eukaryotic cell.
- 41. The isolated cell of Claim 40 wherein the eukaryotic cell is a mammalian cell such as from a livestock animal (e.g. sheep, pig, horse, goat, llama, cow) or part thereof (e.g. wool, leather).
- 42. The isolated cell of Claim 40 wherein the eukaryotic cell is a non-mammalian animal cell (e.g. avian species such as ostrichs, emus, ducks, chickens, turkeys).
- 43. The isolated cell of Claim 40 wherein the eukaryotic cell is a plant cell.
- 44. The isolated plant cell of Claim 43 wherein the cell is part of a plant callus or a whole plant.
- 45. The isolated plant cell of Claim 44 wherein the whole plant is an ornamental or flowering plant or a part thereof.
- 46. The isolated plant cell of Claim 45 wherein the plant part is a flower, root, leaf, stem, seed, fruit or fiber.
- 47. The isolated plant cell of Claim 45 wherein the plant is selected from a rose, carnation, lisianthus, petunia, lily, tulip, pansy, gerbera or chrysanthemum.
- 48. The isolated cell of Claim 34 wherein the CFM is a GFP or derivative or homolog thereof.

- 171 -

- 49. The isolated cell of Claim 48 wherein the homolog of GFP is a non-fluorescent GFP homolog thereof.
- A plant or part of a plant wherein said plant or plant part comprises cells genetically modified to enable production of a CFM which alone or in combination with one or other molecules imparts an altered visual characteristic to said cells when visualized by a human eye in the absence of excitation by extraneous non-white light or particle emission.
- The plant or part of a plant of Claim 50 wherein the CFM is derived from Anemonia majano, Anemonia sulcata, Clavularia sp, Zoanthus sp, Discosoma sp (e.g. Discosoma striata), Aequorea sp (e.g. Aequorea victoria), Anthozoa sp, Cassiopea sp, (e.g. Cassiopea xamachana), Millepora sp, Acropora sp (e.g. Acropora aspera and Acropora nobilis), Montipora sp, Porites murrayensis, Pocillopora damicormis, Pavona descussaca, Acanthastrea sp, Platygyra sp or Caulastrea sp.
- 52. The plant or part of a plant of Claim 51 wherein the CFM comprises an amino acid sequence in its N-terminal end selected from SVIAK (SEQ ID NO:5), (M)SVIAT (SEQ ID NO:6), SGIAT (SEQ ID NO:7), SVIVT (SEQ ID NO:8) or SVSAT (SEQ ID NO:9).
- The plant or part of a plant of Claim 51 wherein the CFM comprises a polypeptide having an amino acid sequence selected from the list comprising SVIAT QMTY KVYM SGT (SEQ ID NO:10), SVIAT QMTY KVYM PGT (SEQ ID NO:11), SVIAT QVTY KVYM SGT (SEQ ID NO:12), SGIAT QMTY KVYM SGT (SEQ ID NO:13), SVIVT QMTY KVYM SGT (SEQ ID NO:14), SVSAT QMTY KVYM SGT (SEQ ID NO:15), SVIAK QMTY KVNM SGT (SEQ ID NO:16), SVIAK QMTY KVYM SDT (SEQ ID NO:17) and SVIAK QMTY X<sub>1</sub>X<sub>2</sub>YX<sub>3</sub> SGT (SEQ ID NO:18) wherein X<sub>1</sub>, X<sub>2</sub> and X<sub>3</sub> may be any amino acid provided that X<sub>1</sub> is not K; X<sub>2</sub> is not V; X<sub>3</sub> is not M.
- 54. The plant or part of a plant of Claim 53 wherein the CFM comprises a

- 172 -

polypeptide having an amino acid sequence selected from the list comprising SEQ ID NOs:20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178, 180, 190, 192, 194, 196, 198, 200 and 202 provided that, where the said amino acid sequence comprises the sequence SVIAK QMTY X<sub>1</sub>X<sub>2</sub>YX<sub>3</sub> SGT, X<sub>1</sub> is not lysine, X<sub>2</sub> is not valine, and X<sub>3</sub> is not methionine or an amino acid sequence having at least 60% similarity to any one or more of the above referenced sequences.

- 55. The plant or part of a plant of Claim 50 wherein the whole plant is an ornamental or flowering plant or a part thereof.
- 56. The plant or part of a plant of Claim 55 wherein the plant part is a flower, root, leaf, stem, seed, fruit or fiber.
- 57. The plant or part of a plant of Claim 55 wherein the plant is selected from a rose, carnation, lisianthus, petunia, lily, tulip, pansy, gerbera or chrysanthemum.
- 58. The plant or part of a plant of Claim 50 wherein the CFM is a GFP or derivative or homolog thereof.
- 59. The plant or part of a plant of Claim 58 wherein the homolog of GFP is a non-fluorescent GFP.
- 60. A cut flower from a plant of any one of Claims 50 to 59.
- 61. An extract from a plant or part of a plant of any one of Claims 50 to 59.
- 62. The extract of Claim 61 wherein the extract is a flavoring or food additive, beverage or juice, or coloring agent.

- 173 -

- 63. Isolated hemp material from a plant of any one of Claims 50 to 54, 58 or 59.
- 64. Cotton from a plant of any one of Claims 50 to 54, 58 or 59.
- 65. A composition comprising a CFM of any one of Claims 18 to 33.
- A method for generating a transgenic plant or part of a plant, wherein said 66. plant or plant part comprises cells genetically modified to enable production of a CFM which alone or in combination with one or other molecules imparts an altered visual characteristic to said cells when visualized by a human eye in the absence of excitation by extraneous non-white light or particle emission, said method comprising introducing into said cells an isolated nucleic acid molecule comprising a nucleotide sequence selected from the list comprising SEQ ID NOs:19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, 131, 133, 135, 137, 139, 141, 143, 145, 147, 149, 151, 153, 155, 157, 159, 161, 163, 165, 167, 169, 171, 173, 175, 177, 179, 189, 191, 193, 195, 197, 199 and 201 or a nucleotide sequence having at least 60% similarity to one or more of the above referenced sequences or a nucleotide sequence capable of hybridizing to one of the above referenced sequences or a complementary form thereof under low stringency conditions and regenerating a transgenic plant therefrom.
- 67. The method of Claim 66 wherein said CFM is derived from Anemonia majano, Anemonia sulcata, Clavularia sp, Zoanthus sp, Discosoma sp (e.g. Discosoma striata), Aequorea sp (e.g. Aequorea victoria), Anthozoa sp, Cassiopea sp, (e.g. Cassiopea xamachana), Millepora sp, Acropora sp (e.g. Acropora aspera and Acropora nobilis), Montipora sp, Porites murrayensis, Pocillopora damicormis, Pavona descussaca, Acanthastrea sp, Platygyra sp or Caulastrea sp.
- · 68. The method of Claim 67 wherein the CFM comprises an amino acid

- 174 -

sequence in its N-terminal end selected from SVIAK (SEQ ID NO:5), (M)SVIAT (SEQ ID NO:6), SGIAT (SEQ ID NO:7), SVIVT (SEQ ID NO:8) or SVSAT (SEQ ID NO:9).

- The method of Claim 67 wherein the CFM comprises a polypeptide having an amino acid sequence selected from the list comprising SVIAT QMTY KVYM SGT (SEQ ID NO:10), SVIAT QMTY KVYM PGT (SEQ ID NO:11), SVIAT QVTY KVYM SGT (SEQ ID NO:12), SGIAT QMTY KVYM SGT (SEQ ID NO:13), SVIVT QMTY KVYM SGT (SEQ ID NO:14), SVSAT QMTY KVYM SGT (SEQ ID NO:15), SVIAK QMTY KVNM SGT (SEQ ID NO:16), SVIAK QMTY KVYM SDT (SEQ ID NO:17) and SVIAK QMTY X<sub>1</sub>X<sub>2</sub>YX<sub>3</sub> SGT (SEQ ID NO:18) wherein X<sub>1</sub>, X<sub>2</sub> and X<sub>3</sub> may be any amino acid provided that X<sub>1</sub> is not K; X<sub>2</sub> is not V; X<sub>3</sub> is not M.
- The method of Claim 69 wherein the CFM comprises a polypeptide having an amino acid sequence selected from the list comprising SEQ ID NOs:20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178, 180, 190, 192, 194, 196, 198, 200 and 202 provided that, where the said amino acid sequence comprises the sequence SVIAK QMTY  $X_1X_2YX_3$  SGT,  $X_1$  is not lysine,  $X_2$  is not valine, and  $X_3$  is not methionine or an amino acid sequence having at least 60% similarity to any one or more of the above referenced sequences.
- 71. The method of Claim 66 wherein the plant part is plant callus.
- 72. The method of Claim 66 wherein the plant part is a flower, root, leaf, stem, seed, fruit or fiber.
- 73. The method of Claim 66 wherein the plant is an ornamental or flowering plant or a part thereof.

WO 02/070703

- 175 -

PCT/GB02/00928

- 74. The method of Claim 73 wherein the plant is selected from a rose, carnation, lisianthus, petunia, lily, tulip, pansy, gerbera or chrysanthemum.
- 75. The method of Claim 66 wherein the CFM is a GFP or derivative or homolog thereof.
- 76. The method of Claim 75 wherein the homolog of GFP is a non-fluorescent GFP homolog thereof.
- An isolated antibody specific for a CFM, said CFM comprising an amino acid sequence in its N-terminal end selected from SVIAK (SEQ ID NO:5), (M)SVIAT (SEQ ID NO:6), SGIAT (SEQ ID NO:7), SVIVT (SEQ ID NO:8) or SVSAT (SEQ ID NO:9).
- The isolated antibody of Claim 77 wherein the CFM comprises a polypeptide having an amino acid sequence selected from the list comprising SVIAT QMTY KVYM SGT (SEQ ID NO:10), SVIAT QMTY KVYM PGT (SEQ ID NO:11), SVIAT QVTY KVYM SGT (SEQ ID NO:12), SGIAT QMTY KVYM SGT (SEQ ID NO:13), SVIVT QMTY KVYM SGT (SEQ ID NO:14), SVSAT QMTY KVYM SGT (SEQ ID NO:15), SVIAK QMTY KVNM SGT (SEQ ID NO:16), SVIAK QMTY KVYM SDT (SEQ ID NO:17) and/or SVIAK QMTY X<sub>1</sub>X<sub>2</sub>YX<sub>3</sub> SGT (SEQ ID NO:18) wherein X<sub>1</sub>, X<sub>2</sub> and X<sub>3</sub> may be any amino acid provided that X<sub>1</sub> is not K; X<sub>2</sub> is not V; X<sub>3</sub> is not M.
- 79. The isolated antibody of Claim 77 or 78, wherein the CFM comprises a polypeptide having an amino acid sequence selected from the list comprising SEQ ID NOs:20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178, 180, 190, 192, 194, 196, 198, 200 and/or 202 or an amino acid sequence having at least 60% similarity to any one or more of the above referenced sequences.

- 176 -

- 80. A biomatrix comprising a CFM, said CFM comprising a polypeptide which, in a cell, alone or together with one or more other molecules imparts an altered visual characteristic to said cell when visualized by a human eye in the absence of excitation by extraneous non-white light or particle emission.
- 81. The biomatrix of Claim 80 wherein the CFM is derived from Anemonia majano, Anemonia sulcata, Clavularia sp, Zoanthus sp, Discosoma sp (e.g. Discosoma striata), Aequorea sp (e.g. Aequorea victoria), Anthozoa sp, Cassiopea sp, (e.g. Cassiopea xamachana), Millepora sp, Acropora sp (e.g. Acropora aspera and Acropora nobilis), Montipora sp, Porites murrayensis, Pocillopora damicormis, Pavona descussaca, Acanthastrea sp, Platygyra sp or Caulastrea sp.
- 82. The biomatrix of Claim 81 wherein the CFM comprises an amino acid sequence in its N-terminal end selected from SVIAK (SEQ ID NO:5), (M)SVIAT (SEQ ID NO:6), SGIAT (SEQ ID NO:7), SVIVT (SEQ ID NO:8) or SVSAT (SEQ ID NO:9).
- The biomatrix of Claim 82 wherein the CFM comprises a polypeptide having an amino acid sequence selected from the list comprising SVIAT QMTY KVYM SGT (SEQ ID NO:10), SVIAT QMTY KVYM PGT (SEQ ID NO:11), SVIAT QVTY KVYM SGT (SEQ ID NO:12), SGIAT QMTY KVYM SGT (SEQ ID NO:13), SVIVT QMTY KVYM SGT (SEQ ID NO:14), SVSAT QMTY KVYM SGT (SEQ ID NO:15), SVIAK QMTY KVNM SGT (SEQ ID NO:16), SVIAK QMTY KVYM SDT (SEQ ID NO:17) and SVIAK QMTY X<sub>1</sub>X<sub>2</sub>YX<sub>3</sub> SGT (SEQ ID NO:18) wherein X<sub>1</sub>, X<sub>2</sub> and X<sub>3</sub> may be any amino acid provided that X<sub>1</sub> is not K; X<sub>2</sub> is not V; X<sub>3</sub> is not M.
- 84. The biomatrix of Claim 83 wherein the CFM comprises a polypeptide having an amino acid sequence selected from the list comprising SEQ ID NOs:20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150,

- 177 -

152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178, 180, 190, 192, 194, 196, 198, 200 and 202 provided that, where the said amino acid sequence comprises the sequence SVIAK QMTY  $X_1X_2YX_3$  SGT,  $X_1$  is not lysine,  $X_2$  is not valine, and  $X_3$  is not methionine or an amino acid sequence having at least 60% similarity to any one or more of the above referenced sequences.

- 85. The biomatrix of Claim 80 wherein the cell is a prokaryotic cell.
- 86. The biomatrix of Claim 80 wherein the cell is a eukaryotic cell.
- 87. The biomatrix of Claim 86 wherein the eukaryotic cell is a mammalian animal cell.
- 88. The biomatrix of Claim 86 wherein the eukaryotic cell is a non-mammalian animal cell.
- 89. The biomatrix of Claim 88 wherein the non-mammalian animal cell is a plant cell.
- 90. The biomatrix of Claim 89 wherein the plant cell is part of a plant callus or a whole plant.
- 91. The biomatrix of Claim 90 wherein the whole plant is an ornamental or flowering plant or a part thereof.
- 92. The biomatrix of Claim 91 wherein the plant part is a flower, root, leaf, stem, seed, fruit or fiber.
- 93. The biomatrix of Claim 91 wherein the plant is selected from a rose, carnation, lisianthus, petunia, lily, tulip, pansy, gerbera or chrysanthemum.

- 178 -

- 94. The biomatrix of any one of Claims 80 to 93, wherein the CFM is a GFP or derivative or homolog thereof.
- 95. The biomatrix of Claim 94 wherein the homolog of GFP is a non-fluorescent GFP.
- 96. The biomatrix of any one of Claims 80 to 95 wherein the said biomatrix is a sunscreen.
- 97. The biomatrix of any one of Claims 80 to 95 wherein the said biomatrix is a cosmetic,
- 98. The biomatrix of any one of Claims 80 to 95 wherein the said biomatrix is a light-filtering composition.
- 99. The biomatrix of any one of Claims 80 to 95 wherein the said biomatrix is a photon trap.
- 100. The biomatrix of any one of Claims 80 to 95 wherein the said biomatrix is a reporter molecule.
- 101. A diagnostic assay comprising screening for the presence of a CFM wherein the nucleic acid molecule encoding said CFM is expressed in a cell.
- The diagnostic assay of Claim 101 wherein said nucleic acid molecule comprises a nucleotide sequence encoding a CFM comprising a polypeptide having an amino acid sequence in its N-terminal end selected from SVIAK (SEQ ID NO:5), (M)SVIAT (SEQ ID NO:6), SGIAT (SEQ ID NO:7), SVIVT (SEQ ID NO:8) or SVSAT (SEQ ID NO:9).
- 103. The diagnostic assay of Claim 102 wherein said nucleic acid molecule

- 179 -

comprises a nucleotide sequence encoding a CFM comprising a polypeptide having an amino acid sequence selected from the list comprising SVIAT QMTY KVYM SGT (SEQ ID NO:10), SVIAT QMTY KVYM PGT (SEQ ID NO:11), SVIAT QVTY KVYM SGT (SEQ ID NO:12), SGIAT QMTY KVYM SGT (SEQ ID NO:13), SVIVT QMTY KVYM SGT (SEQ ID NO:14), SVSAT QMTY KVYM SGT (SEQ ID NO:15), SVIAK QMTY KVNM SGT (SEQ ID NO:16), SVIAK QMTY KVYM SDT (SEQ ID NO:17) and SVIAK QMTY X<sub>1</sub>X<sub>2</sub>YX<sub>3</sub> SGT (SEQ ID NO:18) wherein X<sub>1</sub>, X<sub>2</sub> and X<sub>3</sub> may be any amino acid provided that X<sub>1</sub> is not K; X<sub>2</sub> is not V; X<sub>3</sub> is not M.

The diagnostic assay of Claim 103 wherein said nucleic acid molecule comprises a nucleotide sequence encoding a CFM comprising the CFM comprises a polypeptide having an amino acid sequence selected from the list comprising SEQ ID NOs:20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178, 180, 190, 192, 194, 196, 198, 200 and 202 provided that, where the said amino acid sequence comprises the sequence SVIAK QMTY X<sub>1</sub>X<sub>2</sub>YX<sub>3</sub> SGT, X<sub>1</sub> is not lysine, X<sub>2</sub> is not valine, and X<sub>3</sub> is not methionine or an amino acid sequence having at least 60% similarity to any one or more of the above referenced sequences.

105. The diagnostic assay of any one of Claims 101 to 104 wherein said nucleic acid molecule comprises a nucleotide sequence selected from the list comprising SEQ ID NOs:19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, 131, 133, 135, 137, 139, 141, 143, 145, 147, 149, 151, 153, 155, 157, 159, 161, 163, 165, 167, 169, 171, 173, 175, 177, 179, 189, 191, 193, 195, 197, 199 and 201 or a nucleotide sequence having at least 60% similarity to one or more of the above referenced sequences or a nucleotide sequence capable of hybridizing to one of the above referenced sequences or a complementary form thereof under low stringency conditions.

- 180 -

PCT/GB02/00928

WO 02/070703

The diagnostic assay of Claim 101 wherein the CFM comprises an amino acid sequence in its N-terminal end selected from SVIAK (SEQ ID NO:5), (M)SVIAT (SEQ ID NO:6), SGIAT (SEQ ID NO:7), SVIVT (SEQ ID NO:8) or SVSAT (SEQ ID NO:9).

- The diagnostic assay of Claim 101 wherein the CFM comprises a polypeptide having an amino acid sequence selected from the list comprising SVIAT QMTY KVYM SGT (SEQ ID NO:10), SVIAT QMTY KVYM PGT (SEQ ID NO:11), SVIAT QVTY KVYM SGT (SEQ ID NO:12), SGIAT QMTY KVYM SGT (SEQ ID NO:13), SVIVT QMTY KVYM SGT (SEQ ID NO:14), SVSAT QMTY KVYM SGT (SEQ ID NO:15), SVIAK QMTY KVNM SGT (SEQ ID NO:16), SVIAK QMTY KVYM SDT (SEQ ID NO:17) and SVIAK QMTY X<sub>1</sub>X<sub>2</sub>YX<sub>3</sub> SGT (SEQ ID NO:18) wherein X<sub>1</sub>, X<sub>2</sub> and X<sub>3</sub> may be any amino acid provided that X<sub>1</sub> is not K; X<sub>2</sub> is not V; X<sub>3</sub> is not M.
- The diagnostic assay of Claim 107 wherein the CFM comprises a polypeptide having an amino acid sequence selected from the list comprising SEQ ID NOs:20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178, 180, 190, 192, 194, 196, 198, 200 and 202 provided that, where the said amino acid sequence comprises the sequence SVIAK QMTY  $X_1X_2YX_3$  SGT,  $X_1$  is not lysine,  $X_2$  is not valine, and  $X_3$  is not methionine or an amino acid sequence having at least 60% similarity to any one or more of the above referenced sequences.
- Use of a nucleic acid molecule encoding a CFM, said CFM, in a cell, alone or together with one or more other molecules imparts an altered visual characteristic to said cell when visualized by a human eye in the absence of excitation by extraneous non-white light or particle emission, in the manufacture of a cell which produces said CFM.

WO 02/070703 PCT/GB02/00928

- 181 -

- 110. Use of the CFM of Claim 109 wherein the cell is a prokaryotic cell.
- 111. Use of the CFM of Claim 109 wherein the cell is a eukaryotic cell.
- 112. Use of the CFM according to Claim 111 wherein the eukaryotic cell is a mammalian animal cell.
- 113. Use of the CFM according to Claim 111 wherein the eukaryotic cell is a non-mammalian cell.
- 114. Use of the CFM according to Claim 113, wherein the non-mammalian cell is a plant cell.

Figure 1

### SVIAKOMTYKVYMSDTVNGHYFEVEGDGKGKPYEGEQTVKLTVTKGGPLPFAWDILSPQS SVIAKOMTYKVYMSGTVNGHYFEVEGDGKGKPYEGEQTVRLTVTKGGPLPFAWDILSPQS SVIAKQMIYKVYMSGTVNGHYFEVEGDGKGKPYEGEQTVKLTVTKGGPLPFAWDILSPQS SVIAKQMIYKVYMSGTVNGHYFEVEGDGKGKPYEGEQTVKLTVTKGGPLPFAWDILSPRC SVIAKOMIYKVYMSGTVNGHYFEVEGDGKGKPYEGEQTVRLTVTKGGPLPFAWDILSPQS SVIAKOMTYKVYMSGTVNGHYFEVEGDGKGKPYEGEQTVKLTVTKGGPLPFAWDILSPQC SVIAKQMTYKVYMSGTVNGHYFEVEGDGKGKPYEGEQTVRLTVTKGGPLPFAWDILSPQS SVIAKOMTYKVYMSGTVNGHYFEVEGDGKGKPYEGEQTVKLTVTXGGPLPFAWDIXSPQS SVIAKQMTYKVYMSGTVNGHYFEVEGDGKGKPYEGEQTVKLTVTKGGPLPFAWDILSPQC SVIAKQMTY: "Y 'SGTVXGHYFEVEGDGKGKPYEGEQTVKLTVTKGGPLPFAWDILSPQC SVIAKQMTYKVYMSGTVNGHYFEVEGDGKGKPYEGEQTVKLTVTKGGPLPFAWDILSPQS SVIAKQMTYKVYMSGTVNGHYFEVEGDGKGKPYEGEQTVKLTVTKGGPLPFAWDILSPQS SVIAKQMTYKVYMSGTVNGHYFEVEGDGKGKPYEGEQTVKLTVTKGGPLPFAWDILSPQC SVIAKQMTYKVYMSGTVNGHYFEVEGDGKGKPYEGEQTVKLTVTKGGPLPFAWDILSPQS SVIAKQMTYKVYMSGTVNGHYFEVEGDGKGKPYEGEQTVKLTVTKGGPLPFAWDILSPQS SVIAKOMTYKVYMSGTVNGHYFEVEGDGKGKPYEGEQTVKLTVTEGGPLPFAWDILSPQS SVIAKQMTYKVYMSDTVNGHYFEVEGDGKGKPYEGEQTVKLTVTKGGPLPFAWDILSPQS SVIAKOMTYKVYMSGTVNGHYFEVEGDGKGKPYEGEQTVRLTVTKGGPLPFAWDILSPQS SVIAKOMTYKVYMSGTVNGHYFEVEGDGKGKPYEGEQTVRLTVTKGGPLPFAWDILSPQS SVIAKQMTYKVYMSGTVNGHYFEVEGDGKGRPYEGEQTVKLTVTKGGPLPFAWDILSPQS SVIAKOMTYKVYMSGTVNGHYFXVEGDGKGKPYEGEQTVKLTVTKGGPLPFAWDILSPQS SVIAKOMTYKVYMSGTVNGHYFEVQGDGKGKPYEGEQTVKJTVTKGGPLPFAWDILSPQS SVIAKOMTYKVYMSGTVNGHYFEVEGDRKGKPYEGEQTVKLTVTKGGPLPFAWDILSPQC SVIAKOMTYKVYMSGTVNGHYFEVEGDGKGKPYEGEQTVKLTVTKGGPLPFAWDILSPQC SVIAKOMTYKVYMSGTVNGHYFEVEGDGKGKPYEGEQTVKLTVTKGGPLPFAWDILSPQC SVIAKOMTYKVYMSGTVNGHYFEVEGDGKGKPYEGEQTVKLTVTKGGPLPFAWDILSPRC SVIAKOMTYKVYMSGTVNGHYFEVEGDGKGKPYEGEQTVKLTVTKGGPLPFAWDILSPQS SVIAKOMTYKVYMSGTVNGHYFEVEGDGKGKPYEGEQTVKLTVTKGGPLPFAWDILSPOS SVIAKQMIYKVNMSGTVNGHYFEVEGDGKGKPYEGEQTVKLTVTEGGPLPFAWDILSPQS SVIAKOMIYKVYMSGTVNGHYFEVEGDGKGKPYEGEQTVKLTVTKGGPLPFAWDILSPOC SVIAKOMTYKVYMSGTVNGHYFEVEGDGKGKPYEGEQTVKLTVTEGGPLPFAWDILSPQS SVIAKQMIYKVYMSGTVNGHYFEAEGDGKGKPYEGBQTVKLTVTKGGPLPFAWDILSPQS SVIAKOMTYKVYMSGTVNGHYFEVEGDGKGKPYEGEQTVKLTVTKGGPLPFAWDILSPQS SVIAKOMTYKVYMSGTVNGHYFEVEGDGKGKPYEGEQTVKLTVTKGGPLPFAWDILSPQS , 如果水水水水水水水水水水 不不不不 不不不不不不不不 不 不 不 不 人名英格兰人姓氏 不 PM1Asv-rep.pep PM1Csv-rep.pep GPd58-2sv.pep Ce61-3sv.pep Ce61-4sv.pep Ce61-5sv.pep Ce61-7sv.pep Acasv-D.pep Acasv-A.pep LGAsv-A.pep LGASV-C.pep Acagv-C.pep LGAsv-D.pep LGASV-E.pep Aasv-3.pep Aasv-P.pep Pavsv-A.pep Pavsv-C.pep Pavsv-B.pep Misv-A.pep Misv-B.pep Misv-F.pep PMsv-5.pep PPsv-1.pep PPsv-2.pep PPsv-3.pep PPsv-4.pep PPsv-5.pep PMsv-4.pep PPsv-6.pep RIsv-1.pep RTsv-2.pep RTsv-3.pep NO:26] NO:38] NO:46] NO:28 NO:30 NO:32 NO:34 NO:36 NO:40 NO:42 NO:44 NO:48 NO:50 NO:52] NO:54] NO:56 NO:58 NO: 60 NO:62] NO:66] NO:72] NO:74] NO:64 NO:68 NO:70 NO:76 NO:78] NO:80 NO:86 88888 88888 A П H A Н H H B ü Π H Ü H H 88 H A H H 品品 SEQ Oas] Oas] Oas] SEQ SEQ SEQ SEQ SEO SEQ SEO SEQ SEO SEQ SEQ SEQ SEQ SEQ SEO SEO

TVSNDSSIQGNCFTYHVKFS TVSNDSSIQGNCFIYNVKFS

TVSNDSSIQGNCFIYNVKFS
TVSNDSSIQGNCFIYNVKFS
TVSNDSSIQGNCFIYNVKFS
TVSNDSSIQGNCFIYNVKFS
TVSNDSSIQGNCFIYHVKFS
TVSNDSSIQGNCFIYHVKFS
TVSNDSSIQGNCFIYNVKFS

TYSNDSSIQGNCFIYNVKFS
TYSNDSSIQGNCFIYHVKFS
TYSNDSSIQGNCFIYHVKFS
TYSNDSSIQGNCFIYHVKFS
TYSNDSSIQGNCFIYHVKFS
TYSNDSSIQGNCFIYHVKFS
TYSNDSSIQGNCFIYHVKFS
TYSNDSSIQGNCFIYHVKFS
TYSNDSSIQGNCFIYHVKFS
TYSNDSSIQGNCFIYHVKFS

\*\*\*\* \* \*\* \*\* \*\*\*\*\*\*\*\*\*

TVSNDSSIQGNCFTYHVKFS
TVSNDSSIQGNCFTYHVKFS

TVSNDSSIQGNCFIYNVKFS
TVSNDSSIQGNCFIYNVKFS
TVSNDSSIQGNCFIYHVKFS
TVSNDSSIQGNCFIYHVKFS
TVSNDSSIQGNCFIYNVKFS
TVSNDSSIQGNCFIYNVKFS
TVSNDSSIQGNCFIYNVKFS
TVSNDSSIQGNCFIYNVKFS

GLNFPPNC	-3.pe	[SEQ ID NO:86]
GLNFPPNC	RTsv-2.pep	[SEQ ID NO:84]
GLNFPPNC	RTsv-1.pep	E0 10
GLNFPPNC		ID NO:8
GLNFPPN	Pavsv-B.pep	ID NO:7
GLNFPPN	Pavsv-A.pep	ID NO:7
GLNFPPN	PPsv-6.pep	TD NO:7
GLNFPPN	PPsv-5.pep	A
GLNPPPN	PPsv-4.pep	ID NO:
GLNFPPN	PPsv-3.pep	ID NO:6
GLNFPPN	PPsv-2.pep	ID NO:6
GLNFPPN	SV-1.	9:0X GI
GLNFPPN		A
GLNFPPN		ID NO:
GLNFPPN	PM1Csv-rep.pep	A
GLNFPPN	PMlAsv-rep.pep	A
GLNFPPN	Misv-F.pep	[SEQ ID NO:54]
GLNPPN	Misv-B.pep	[SEQ ID NO:52]
GENPPP	Misv-A.pep	(SEQ ID NO:50]
GLNFPPN	•	
GLNFPPN	LGAsv-D.pep	
GLDFPPN	LGAsv-C.pep	[SEQ ID NO:44]
GLNFPPN	LGASV-A. pep	
GLNFPPN	GPd58-2sv.pep	
GLNEPPN	Ce61-7sv.pep	[SEQ ID NO:38]
GLNFPPN		[SEQ ID NO:36]
GT,NFD	Ce61-4sv.pep	[SEQ ID NO:34]
GLNFPPN	Cecl-3sv.pep	[SEQ ID NO:32]
GLNFPPN	ė.	[SEQ ID NO:30]
GLNFPPN	Acasy-C.pep	[SEQ ID NO:28]
GLNFPPN	Acasv-A.pep	a
GLNPPP	Aasv-P.pep	[SEQ ID NO:24]
GLNFPPN	m	[SEQ ID NO:22]
GLNFPP	Aasv-1.pep	A

NGPVMQKKTQGWEDNTERLYARDGMLIGNNFWALKLEGGQRSL\*-----NGPVMQKKTQGWEPNTGRLFARDGMLIGNNFMALKLEGGGHYLCEFKSTYKAKK ngpvmokktogenepnterllardgmlignnfmalkleggghylcefkstykark IGPVMQKKTQGWEPNTERLSARDGMLIGNNFMALKLEGGGHYLCEFKSTYKARK ngpvmokktrgwephserlfarggmlignnfmalkleggghylcgfkttykkakk ngpvmokktogwephserlfarggmlignnfmapkleggchylcefkttykakk igpvmokktogwepnterlsardgmlignnfmalkleggghylcefkstykark IGPVMQKKTQGWEPHSERLFARDGMLIGNNFMALKXEGGGXYLCEFKSTYKAKK igpvmokktogwephserlfarganlignnfmalkleggghylcefktitykakk 1GPVMQKKTQGWBPHSERLFARGGMLIGNNFMALKLEGGGHYLCEFKTTYKAKK IGPVMQKKTQGWEPNTERLFARDGMLIGNNFMALKLEGGGHYLCEFKSTYKAKK igpvmokktogwepnterlfardgmlignnfmalkleggghylcefkstykakk IGPVMQKKTQGWEPHSERLFARGGMLIGNNFMALKLEGGGHYLCEFKTTYKAKK |GPVMQKKTQGWEDNTERLFARDGMLIGNNFMALKLEGGGHYLCEFKSTYKAKK |GPVMQKKTQGWEDNTERLFARDGML||GNNFMALKLEGGGHYLCEFKSTYKAKK GPVMQKKTQGWEPNTERLFARDGMLIGNNFMALKLEGGGHYLCEFKSTYKARK |GPVMQKKTQGWEPNTERLFARDGMLIGNNFMALKLEGGGHYLCEFKSTYKARK |GPVMQKKTQGWEDNTERLFARDGML||GNNFNALKLEGGGHYLCEFKSTYKARK |GPVMQKKTQGWEPNTGRLFARDGMLIGNNFMALKLEGGGHYLCEFKSTYKAKK GPVMQKKTQGWEPNTERLYARDGMLIGNNFMALKLEGSGHYTCEFKSTYKAKK gpvmokktogwepnterlfardgvlignnfmalkleggghylcefkstykakk gpumokktogwepnterlyardgmlignnfwalkleggorsl\*----gpvmokktiogwedhserlfarggmlignnfmalkleggghylcgfkttykakk gpumokktogwephserlfarggmlignnfmalkleggghylcgfkttykakk gpumokktogwephserlfarggmlignnfmalkleggghylcgfkttykakk gpvmokktrgwephserlfarggmlignnfmalkleggghylcgfkttykakk gPvmokktogwepnterlfardgmlignnfmalkleggghylcefkstykakk gpumokktogwepnterlfardgmlignnfmalkleggghylcefkstykakk gpvmokktogwv pnterl fardgml i gnnfmalkleggghyl cefkstykakk GPVMQKKTQGWEPHSERLFARGGMLIGNNFMAPKLEGGGHYLCEPKTTYKAKK GPVMQKKTQGWEPNTERLFARDGMLIGNNFMALKLEGGGHYLCEFKSTYKAKK gpumokktogwepnterlsardgmlignnfmalkleggghylcefkstykark 3PVMQKKTQGWEPNTERLFARDGMLIGNNFMALKLEGGGHYLCEFKSTYKAKK gpumokktogwepnterlfardgmlignnfmalkleggghylcefkstykakk 化妆妆器 化水子聚聚苯甲基 宋 宋宋 秋秋 " 女 水火,水水水水的石块水水水水水

PVMMPGYHYVDRKLDVTNHNKDYTSVEQCEISIARKPVVA	PVKWPGYHYVDRKT, DVTWHNKTOYT SVEOCET STAD V DAY	PVKMPGYHCVDRKIDVINHNKOYTSVEOREISIARKPVVA	PVKMPGYHYVDRKLDVTNHNKDYTSVEQCEISIARKPVVA	PVKMPGYHYVDRKLDVTNHNKDYTSVEOREISIARKPVVA	PVMMPGYHYVDRKLDVTNHNKDYTSVEOCEISLARKPVVA	1 3 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	PVKMPGYHYVDRKLDVTNHNKDYTSVEOCEISIARKPVVA	PVMMPGYHYVDRKLDVTNHNKOYTSVEOCEISIARKPVVA	PVMMPGYHYVDRKIDVTNHNKDYTSVEOCEISIARKPVVA	PVKWPGYHYVDRKLDVTNHNKDYTSVEOCETSTARKPVVA	PVMMPGYHYVDRKIDVTNHNKDYTSVEOCETSTARKPVVA	PVMMPGYHYVDRKLDVTMHNKOYTSVEOCETSTARKOVM	PVKMPGYHYVDRKLDVTNHNKDVTSVEORFISIATSTARKOVVA	PVKMPGYHYVDRKLDVTNHNKDVTSVEORETSTATARKDVIA	PVKMPGYHYVDRKCLDVINENKDYTSVEOREISIARRPVVA	PVMMPGYHYVDRKLDVTNHNKDYTSVEOCEISIARRYVA	PVMMPGYHYVDRKLDVTNHNKDYTSVEOCETSTARKDV773		VWMPGYHYVDRKLDVTNHNKDYTSVEOCETSTARRYDYKVTXVLDVA	PVKMPGYHYVDRKLDVTMHNKDYTSVEOCETSTAPKPMM	PVKMPGYHYVDRKLDVTNHNKDYTSVEOCETSTARVKVVA	PVKMPGYHYVDRKLDVTNHNKDYISVEOCETSIARKPVVA	PVKMPGYHYVDRKLDVTNHNKDYISVEOCEISIARKPVVA	PVMMPGYHYVDRKLDVTNHNKDYTSVEOCETSTARKDVVVA	PVMMPGYHYVDRKLDVTNHNKOYTSVEOCGTSTARKPVN/A	PVMMPGYHYVDRKLDVTNHNKOYTSVEOCETSTARKPVVVA	PVKMPGYHYVDRKLDVTNHNKDYTSVEOCETSTARKPVVA	PVKMPGYHYVDRKLDVTMHNKOYTSVEOREISTARKPVVA	PVMMPGYHYVDRKLDVTNHNKOYTSVEOCETPTARKPVVA	PVMMPGYHYVDRKLDVTNHNKOYTSVEOCEISIARKPVVA	
Aasv-1.pep Aasv-3.pep	Acasv-A.pep	Acasv-C.pep	Acasv-D.pep	Ce61-3sv.pep	Ce61-4sv.pep	Ce61-5sv.pep	Ce61-7sv.pep	GPd58-2sv.pep	LGASV-A.pep	LGASV-C.pep	LGASV-D.pep	LGASV-E.pep	Misv-A.pep	Misv-B.pep	Misv-F.pep	PMIASV-rep.pep	PM1Csv-rep.pep	PMsv-4.pep	PMsv-5.pep	PPsv-1.pep	PPsv-2.pep	PPsv-3.pep	PPsv-4.pep	PPSV-5.pep	Ppev-6.pep	Pavsv-A.pep	Pavsv-B.pep	Pavsv-C.pep	RTsv-1.pep	RTsv-2.pep	RTsv-3.pep
[SEQ ID NO:20] [SEQ ID NO:22] [SEO ID NO:24]	A	A	A I	A	H	A	A	A	A	[SEQ ID NO:44]	A					H		[SEQ ID NO: 60]	A	H	A	A	A	[SEQ ID NO:72]	A	A	[SEQ ID NO:78]	[SEQ ID NO:80]	[SEQ ID NO:82]	[SEQ ID NO:84]	[SEQ ID NO:86]

TCCGTTATCGCTAAACAGATGACCTACAAGTTTATATGTCAGGCACGGTCAATGGACAC TCCGTTATCGCTAAACAGATGACCTACAAGGTTTATATGTCAGGCACGGTCAATGGACAC CCGITATCGCTAAACAGATGACCTACAAGGTTTATATGTCAGACACGGTCAATGGACAC TCCGTTATCGCTAAACAGATGACCTACAAGTTTATATGTCAGGCACGGTCAATGGACAC TCCGTTATCGCTAAACAGATGACCTACAAAGTTTATATGTCAGGCACGGTCAATGGACAC TCCGTTATCGCTAAACAGATGACCTACAAGGTTTATATGTCAGGCACGGTCAATGGACAC TCCGTTATCGCTAAACAGATGACCTACAAAGTTTATATGTCAGGCACGGTCAATGGACAC tccgttatcgctaaacagatgacctacaaggttatatgtcaggcacggtcaatggacac ICCGTTATCGCTAAACAGATGACCTACAAGGTTTATATGTCAGGCACGGTCAATGGACAC ICCGTTAICGCTAAACAGATGACCTACATNGNTTAINTGICAGGCACNGTCNAIGGACAC TCCGTTATCGCTAAACAGATGACCTACAAGGTTTATATGTCAGGCACGGTCAATGGACAC ICCGTTATCGCTAAACAGATGACCTACAAGGTTTATATGTCAGGCACGGTCAATGGACAC TCCGTTATCGCTAAACAGATGACCTACAAGGTTTATATGTCAGGCACGGTCAATGGACAC TCCGTTATCGCTAAACAGATGACCTACAAGGTTTATATGTCAGGCACGGTCAATGGACAC TCCGTTATCGCTAAACAGATGACCTACAAGGTTTATATGTCAGGCACGGTCAATGGACAC TCCGTTATCGCTAAACAGATGACCTACAAGGTTTATATGTCAGGCACGGTCAATGGACAC TCCGTTATCGCTAAACAGATGACCTACAAAGTTTATATGTCAGGCACGGTCAATGGACAC TCCGTTATCGCTAAACAGATGACCTACAAAGTTTATATGTCAGGCACGGTCAATGGACAC TCCGTTATCGCTAAACAGATGACCTACAAGGTTTATATGTCAGACACGGTCAATGGACAC TCCGTTATCGCTAAACAGATGACCTACAAGGTTTATATGTCAGGCACGGTCAATGGACAC TCCGTTATCGCTAAACAGATGACCTACAAGGTTTATATGTCAGGCACGGTCAATGGACAC ICCGTTAICGCTAAACAGATGACCTACAAGGTTTATATGTCAGGCACGGTCAATGGACAC TCCGTTATCGCTAAACAGATGACCTACAAGGTTTATATGTCAGGCACGGTCAATGGACAC ICCGTTAICGCIAAACAGAIGACCIACAAGGITTIAIAIGICAGGCACGGICAAIGGACAC TCCGTTATCGCTAAACAGATGACCTACAAGGTTTATATGTCAGGCACGGTCAATGGACAC ICCGTTATCGCTAAACAGATGACCTACAAGGTTTATATGTCAGGCACGGTCAATGGACAC ICCGTTATCGCTAAACAGATGACCTACAAGGTTTATATGTCAGGCACGGTCAATGGACAC TCCGTTATCGCTAAACAGATGACCTACAAGGTTTATATGTCAGGCACGGTCAATGGACAC ICCGTTATCGCTAAACAGATGACCTACAAGGTTAATATGTCAGGCACGGTCAATGGACAC ICCGTTATCGCTAAACAGATGACCTACAAGGTTTATATGTCAGGCACGGTCAATGGACAC ICCGTTATCGCTAAACAGATGACCTACAAGGTTTATATGTCAGGCACGGTCAATGGACAC rccgttatcgctaaacagatgacctacaaggtttatatgtcaggcacggtcaatggacac rccgttatcgctaaacagatgacctacaaggtttatatgtcaggcacggtcaatggacac **ICCGTTATCGCTAAACAGATGACCTACAAGGTTTATATGTCAGGCACGGTCAATGGACAC** 水子水水水水水 水水水 水水水 大量水水水水 安大 女 水 \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Aasv-1 Aasv-3 Aasv-P Acasv-A Acasv-C Acasv-D	Ce61-3sv Ce61-4ev Ce61-5sv-rep Ce61-7sv-rep GPd58-2sv LGAsv-A	LGASV-C LGASV-D LGASV-E Misv-A Misv-F Misv-F PMIASV-rep	PM1Csv-rep PMsv-4 PMsv-5 PPsv-1 PPsv-3 PPsv-4 PPsv-5	
888888	EQ ID NO:3		[SEQ ID NO:57] [SEQ ID NO:59] [SEQ ID NO:63] [SEQ ID NO:63] [SEQ ID NO:67] [SEQ ID NO:67] [SEQ ID NO:71]	EQ ID NO: EQ ID NO: EQ ID NO: EQ ID NO: EQ ID NO:

GGTTGAAGGCGATGGAAAGGAAAGCCTTACGAGGGGGGAGCAGACGGTAAAG GGTCGAAGGCGATGGAAAAGGAAAGCCTTACGAGGGGGGAGCAGACGGTAAGG GGTCGAAGGCGATGGAAAGGAAAGCCTTACGAGGGGGGAGCAGACGGTAAAG GGTCGAAGGCGAAAAGGAAAGCCTTACGAGGGGGAGCAGACGGTAAGG GGTCGAAGGCGATGGAAAGGAAAGCCTTACGAGGGGGAGCAGACGGTAAAG GGTTGAAGGCGATGGAAAAGGAAAGCCTTACGAGGGGGGAGCAGACGGTAAAG

GGTNGAAGGCGATGGAAAAGGAAAGCCTTACGAGGGGGGAGCAGACGGTAAAG

GGTCGAAGGCGATGGAAAAGGAAAGCCTTACGAGGGGGGGAGCAGACGGTAAGG

GGTCGAAGGCGATGGAAAAGGAAAGCCTTACGAGGGGGGAGCAGCAGTAAAG

GGTTGAAGGCGATGGAAAAGGAAAAGCCTTACGAGGGGGAGCAGCAGTAAAG GGTTGAAGGCGATGGAAAAGGCGTTACGAGGGGGGGGCAGACGGTAAAG GGTCGAAGGCGATGGAAAAGGAAAAGCCTTACGAGGGGGGAGCAGACGGTAAGG SGTTGAAGGCGATGGAAAAGGAAAGCCTTACGAGGGGGAGCAGCAGTAAAG GGTTGAAGGCGATGGAAAAGGAAGGCCTTACGAGGGGGGAGCAGACGGTAAAG SGTTGAAGGCGATGGAAAAGGAAAGCCTTACGAGGGGGGAGCAGAGGTAAAG 

GGTTGAAGGCGATGGAAAAGGAAAGCCTTACGAGGGGGGAGCAGACGGTAAAG

**3GTTGAAGGCGATGGAAAAGGAAAGCCTTACGAGGGGGGAGCAGACGGTAAAG 3GTTGAAGGCGATGGAAAAGGAAAGCCTTACGAGGGGGAGCAGCAGCAGTAAA**G **GTTGAAGGCGATGGAAAAGGAAAGCCTTACGAGGGGGGAGCAGACGGTAAAG** 

**GETTGAAGGCGATGGAAAAGGAAAGCCTTACGAGGGGGGGGAGACGGTAAAG GCTGAAGGCGATGGAAAAGGAAAGCCTTACGAGGGGGAGCAGACGGTAAAG ISTIGAAGGCGAIGGAAAAGGAAAGCCTTACGAGGGGGGGAGCAGACGGTAAAG GTTGAAGGCGATGGAAAAGCCTTACGAGGGGGAGCAGACGTAAA**G 化水水水水水水水水水水水水水水水水水水水水水水水水水水水 计 计计算计算计算计算

TACTITGA	TACTITGA	TACTITGA	TACTTTGA	TACTITGA	TACTTTGA	TACTITGA	TACTITGA	TACTTTGA	TACTITGA	TACTTTGA(	TACTTTGA	TACTTTGA	TACTTIGA	TACTTTGA	TACTTTGA(	TACTTIGA	TACTITGA	TACTITIGAC	TACTTTGAC	TACTTTGAC	TACTTTGAC	TACTTTGAC	TACTITGAC	TACTITGAG	TACTITGAG	TACTTTGAG	TACTTTGAC	TACTTTGAG	TACTTTGAG	TACTTTGAG	TACTITGAG	TACTTTGAG	
Aasv-1 Aasv-3	Aasv-P	Acasv-A	Acasv-C	пť	Ce61-3sv		Ce61-5sv-rep	2	GPd58-2sv	LGABV-A	LGASV-C	LGASV-D	LGASV-E	Misv-A	Misv-B	Misv-F	PM1Asv-rep	PM1Csv-rep	PMsv-4	PMsv-5	PPsv-1	PPsv-2	PPsv-3	PPsv-4	PPsv-5	PPsv-6	Pavsv-A	Pavsv-B	Pavsv-C	RTsv-1	RTsv-2	RTsv-3	
[SEQ ID NO:19] [SEQ ID NO:21]	ID NO:	EQ ID	EQ ID NO:2	EQ ID	EQ ID NO:3	다 다	A	ID NO:3	A	ID NO:	ID NO:	_	E NO	ED NO:	ID NO:5	H NO:5	ID NO:	ID NO:5	io Ro io	ID NO:	ID NO:6	ID NO:6	ID NO:6	ID NO:6	ID NO:7	A	A A	E Q	O O	[SEQ ID NO:81]	EQ ID NO:8	[SEG ID NO:85]	

[SEQ ID NO:19]	Aasv-1	CTCACTGTCACCAAGGGCGGACCTCTGCCATTTTGGTTTGGGATATTCTATCACCACAGAGT
ID NO:2	Aasv-3	CTGACTGTCACCAAGGGCGGACCTCTGCCATTTGCTTGGGATATTTATCACCACAGTCA
[SEQ ID NO:23]	Aasv-p	CTGACTGTCACCAAGGGCGGACCTCTGCCATTTGCTTGGGATATTTTATCACCACTCA
ID NO:2	Acasv-A	CTCACTGTCACCAAGGGCGGACCTCTGCCATTTGGCTTGGGATATTTTATCACCACGGTGT
	Acasv-C.	CTGACTGTCACCAAGGGCGGACCTCTGCCATTTGCTTGGGATATTTATCACCACAGTCA
H	Acasv-D	CICACIGICACCAAGGGCGGACCICTGCCATTIGGGATATITIATCACCACAGIGT
fi	Ce61-3sv	CTGACTGTCACCAAGGGCGGACCTCTGCCATTTTGGGGATATTTTATCACCACAGTCA
ID NO:3	Ce61-4sv	CTCACTGTCACCNAGGGCGGACCTCTGCCATTTGCTTGGGATATTNTATCACCACAGAGT
SON.	Ce61-5sv-rep	CTCACTGTCACCAAGGGCGGACCTCTGCCATTTGGGATATTTTATCACCACAGTGT
A	Ce61-7sv-rep	CTCACTGTCACCAAGGGCGGACCTCTGCCATTTGCTTGGGATATTTATCACCACAGTGT
ED NO:	GPd58-2sv	CTCACTGTCACCAAGGGCGGACCTCTGCCATTTGCTTGGGATATTCTATCACCACAGAGT
_	LGASV-A	CTCACTGTCACCAAGGGCGGACCTCTGCCATTTGCTTGGGATATTCTATCACCACAGAGT
e E	LGASV-C	CICACIGICACCAAGGGGGGACCICIGCCAIIIGGCIIGGGAIAIITIIAICACCACAGIGI
ED NO:	LGASV-D	CICACIGICACCAAGGGGGGACCICTGCCATITGCTTGGGATAITCTATCACCACAGAGT
A	LGASV~E	CICACTGICACCGAGGGCGGACCICTGCCATITGGGGATAITCIATCACCACAGAGT
A	Misv-A	CICACTGICAACGGGGGGAACCICTGCCATTTGGCATATICTATCACAAAGAGT
A	Misv-B	CTGACTGTCACCAAGGGCGGACCTCTGCCATTTGCTTGGGATATTTTATCACCACAGTCA
A	Misv-F	CTGACTGTCACCAAGGGCGGACCTCTGCCATTTGCTTGGGATATTTTATCACCACAGTCA
H	PMLAsv-rep	CICACTGTCACCAAGGGCGGACCTCTGCCATTTGCTTGGGATATTCTATCACCACAGAGT
A	PM1Csv-rep	CICACIGICACCAAGGGCGGACCICTGCCATITGCTTGGGATAITCTAICACCACAGAGT
A	PMsv-4	CTCACTGTCACCAAGGGCGGACCTCTGCCATTTGCTTGGGATATTCTATCACCACAGAGT
ID NO:	PMsv-5	CTCACTGTCACCAAGGGGGGACCTCTGCCATTTTGCTTGGGATATTCTATCACCACAGAGT
Ö.	PPsv-1	CICACTGICACCAAGGGCGGACCICIGCCATITIGCITIGGGAIATITIAICACCACAGIGI
ID NO:	PPsv-2	CICACIGICACCAAGGGCGGACCICIGCCATITGCITGGGAIATITIAICACCACAGIGT
ID NO:	PPsv-3	CTCACTGTCACCAAGGGCGGACCTCTGCCATTTTGCTTTGGGGATATTTTATCACCACAGTGT
A	PPsv-4	CTCACTGTCACCAAGGGCGGGACCTCTGCCATTTGGGGATATTTTATCACCACGGTGT
• •	PPsv-5	CTCACTGTCACCAAGGGCGGACCTCTGCCATTTGCTTGGGATATTCTATCACCACAGAGT
. 7	9-vsgg	CTCACTGTCACCAAGGGCGGACCTCTGCCATTTGCTTGGGATATTCTATCACCACAGAGT
	Pavsv-A	CTCACTGTCACCGAGGGGGGACCTCTGCCATTTGCTTGGGATATTCTATCACACAGAGT
A	Pavsv-B	CTCACTGTCACCAAGGGCGGGACCTCTGCCATTTGCGTTATTTTTTTT
A	Pavsv-C	CTCACTGTCACCGAGGGGGGACCTCTGCCATTTGCTTGGGATATTCTATCACCACAGAGT
A	RTsv-1	CTCACTGTCACCAAGGGGGGACCTCTGCCATTTGCTTGGGATATTCTATCACCACAGAGT
EQ ID NO:8	RTsv-2	CTCACTGTCACCAAGGGGGGACCTCTGCCATTTGCTTGGGATATTCTATCACCACAGAGT
[SEQ ID NO:85]	RTsv-3	CTCACTGTCACCAAGGGCGGACCTCTGCCATTTGCTTGGGATATTCTATCACACAGAGT

					·			Ü		· O	· U	U	U	O	O	U	υ	U	υ	U	ี	ับ	บ	ี่	ี่	ี่	ี่	ฮ	ี่	บิ	ฮ	ฮ	ฮ
Aasv-1		Aasv-P	Acasv-A	Acasv-C	Acasv-D	Ce61-3sv	Ce61-4sv	Ce61-5sv-rep	Ce61-7sv-rep	8-2sv	LGASV-A	LGASV-C	LGASV-D	LGASV-E	Misv-A	Misv-B	Misv-F	PM1Asv-rep	PM1Csv-rep	PMsv-4	PMSV-5	PPsv-1	PPsv-2	PPsv-3	PPsv-4	PPsv-5	PPsv-6	Pavsv-A	Pavsv-B	Pavsv-C	RTsv-1	RTsv-2	RTsv-3
[SEC ON OI DES]	ID NO:2	[SEQ ID NO:23]	(SEQ ID NO:25]	ID NO:2		ID NO:3	A		[SEQ ID NO:37]	[SEQ ID NO:39]	[SEQ ID NO:41]	[SEQ ID NO:43]	[SEQ ID NO:45]	[SEQ ID NO:47]		ID NO:5		ID NO:	H	[SEQ ID NO:59]	[SEQ ID NO:61]	[SEQ ID NO:63]	H	A			A	A	A	o.	[SEQ ID NO:81]	βĄ	[SEQ ID NO:85]

CAGTACGGAAGCATACCATTCACCAAGTACCCTGAAGACATCCCTGACTATGTAAAGCAG CAGTACGGAAGCATACCATTCACCAAGTACCCTGAAGACATCCCTGACTATGTAAAGCAG CAGTACGGAAGCATACCATTCACCAAGTACCCTGAAGACATCCCTGACTATGTAAAGCAG CAGTACGGAAACATACCATTCACCAAGTACCCTGAAGACGTCCCTGACTATGTAAAGCAG CAGTACGGAAGCATACCATTCACCAAGTACCCTGAAGACATCCCTGACTATGTAAAGCAG CAGTACGGAAGCATACCATTCACCAAGTACCCTGAAGACATCCCTGACTATGTAAAGCAG CAGTACGGAAGCATACCATTCACCAAGTACCCTGAAGACATCCCTGACTATGTAAAGCAG CAGTACGGAAGCNTACCATTCACCAAGTACCCTGAAGACATCCCTGACTATGTAAAGCAG CAGTACGGAAGCATACCATTCACCAAGTACCCTGAAGACATCCCTGACTATGTAAAGCAG CAGTACGGAAGCATACCATTCACCAAGTACCCTGAAGACATCCCTGACTATGTAAAGCAG agtacggaagcataccattcaccaagtaccctgaagacatccctgactatgtaaagcag CACTACGGAAGCATACCATTCACCAAGTACCCTGAAGACATCCCTGACTATGTAAAGCAG agtacggaagcataccattcaccaagtaccctgaagacatccctgactatgtaaagcag AGTACGGAAGCATACCATTCACCAAGTACCCTGAAGACATCCCTGACTATGTAAAGCAG agtacggaagcataccattcaccaagtaccctgaagacatccctgactatgtaaagcag agtacggaagcataccattcaccaagtaccctgaagacatccctgactatgtaaagcag 'agtacggaagcataccattcaccaagtaccctgaagacatccctgactatgtaaagcag 'agtacggaagcataccattcaccaagtaccctgaagacatccctgactatgtaaagcag 'AGTACGGAAGCATACCATTCACCAAGTACCCTGAAGACATCCCTGACTATGTAAAGCAG AGTACGGAAGCATACCATTCACCAAGTACCCTGAAGACATCCCTGACTATGTAAAGCAG AGTACGGAAGCATACCATTCACCAAGTACCCTGAGGACATCCCTGACTATGTAAAGCAG agtacggaagcataccattcaccaagtaccctgaagacatccctgactatgtaaagcag 'AGTACGGAAACATACCATTCACCAAGTACCCTGAAGACGTCCCTGACTATGTAAAGCAG AGTACGGAAACATACCATTCACCAAGTACCCTGAAGACGTCCCTGACTATGTAAAGCAG agtacggalacataccaticaccaagtaccctgaagacgtccctgactatgtaaagcag AGTACGGAAACATACCATTCACCAAGTACCCTGAAGACGTCCCTGACTATGTAAAGCAG agtacggaagcataccattcaccaagtaccctgaagacatccctgactatgtaaagcag AGTACGGAAGCATACCATTCACCAAGTACCCTGAAGACATCCCTGACTATGTAAAGCAG AGTACGGAAGCGTACCATTCACCAAGTACCCTGAAGACATCCCTGACTATGTAAAGCAG AGTACGGAAGCATACCATTCACCAAGTACCCTGAAGACATCCCTGACTATGTAAAGCAG AGTACGGAAGCGTACCATTCACCAAGTACCCTGAAGACATCCCTGACTATGTAAAGCAG AGTACGGAAGCATACCATTCACCAAGTACCCTGAAGACATCCCTGACTATGTAAAGCAG AGTACGGAAGCATACCATTCACCAAGTACCCTGAAGACATCCCTGACTATGTAAAGCAG 'agtacggaagcataccattcaccaagtaccctgaagacatccctgactatgtaaagcag 

### **ICATTCCCGGAGGGATATACATGGGAGAGGATCATGAACTTTTGAAGATGGTGCAGTGTGT** ICATICCCGGAGGGATATACATGGGAGAGGATCATGAACTTTTGAAGATGGTGCAGTGTGT TCATTCCCGGAGGATATACATGGGAGAGATCATGAACTTTGAAGATGGTGCAGTGTGT TCATTCCCGGAGGATTTACATGGGAGAGGATCATGAACTTTTGAAGATGGTGCAGTGTGT TCATTCCCGGAGGATTTACATGGGAGAGGATCATGAACTTTGAAGATGGTGCAGTGTGT TCATTCCCGGAGGATATACATGGGAGAGATCATGAACTTTGAAGATGGTGCAGTGTGT TCATTCCCGGAGGATTTACATGGGAGAGAGATCATGAACTTTGAAGATGGTGCAGTGTGT TCATTCCCTGAGGGATATACATGGGAGGGATCATGAACTTCGAAGATGGTGCAGTGTGT TCATTCCCTGAGGGATATACATGGGAGAGGATCATGAACTTCGAAGATGGTGCAGTGTGT TCATTCCCGGAGGGAITTACATGGGAGGATCATGAACTTTGAAGATGGTGCAGTGTGT TCATTCCCGGAGGATTTACATGGGAGAGGATCATGAACTTTGAAGATGGTGCAGTGTGT TCATTCCCTGAGGGATATACATGGGAGAGGATCATGAACTTCGAAGATGGTGCAGTGTGT TCATTCCCTGAGGGATATACATGGGAGAGGATCATGAACTTCGAAGATGGTGCAGTGTGT TCATTCCCTGAGGGATATACATGGGAGAGGATCATGAACTTTGAAGATGGTGCAGTGTGT TCATTCCCGGAGGATATACATGGGAGGATCATGAACTTTTGAAGATGGTGCAGTGTGT ICATTCCCTGAGGGATATACATGGGAGAGATCATGAACTTCGAAGATGGTGCAGTGTGT ICATTCCCTGAGGGATATACATGGGAGAGGATCATGAACTTCGAAGATGGTGCAGTGTG TCATTCCCGGAGGATATACATGGGAGAGGATCATGAACTTTGAAGATGGTGCAGTGTGT TCATTCCCTGAGGGATATACATGGGAGAGATCATGAAGTTTTGAAGATGGTGCAGTATGT ICAITICCCTGAGGGATATACATGGGAGAGGATCGTGAACTTCGAAGATGGTGCAGTGTGT *| Cattccctgaggatatacatgggagaggatcatgaacttcaaagatggtgcagtgtgt* rcaitcccggaggaittacaigggagagagatcaigaactitigaagaiggtgcagigigi Icattcccggaggatttacatgggagagatcatgaactttgaagatggtgcagtgtgt ICATTCCCGGAGGGATTTACATGGGAGAGGATCATGAACTTTTGAAGATGGTGCAGTGTGT CAITCCCTGAGGGAIATACATGGGAGAGGATCATGAACTTCGAAGATGGTGCAGTGTGT Tcattccctgagggatatacatgggagaggatcatgaacttcgaagatggtgcagtgtgt | CATTCCCTGAGGGATATACATGGGAGAGGGATCATGAACTTCGAAGATGGTGCAGTGTGT Icattcccggaggatttacatgggagagatcatgaactttgaagatggtgcagtgtgt Icaitcccggaggatatacatgggagagatcatgaacittgaagatggtgcagtgtgt ICATTCCCTGAGGGATATACATGGAGAGGATCATGAACTTCGAAGATGGTGCAGTGTGT ICATTCCCTGAGGGATATACATGGGAGAGGGATCATGAACTTCGAAGATGGTGCAGTGTGT TCATTCCCTGAGGGATATACATGGGAGAGAGATCATGAACTTCGAAGATGGTGCAGTGTGT Ce61-5sv-rep Ce61-7sv-rep PMlAsv-rep PM1Csv-rep GPd58-2sv Ce61-3sv Ce61-4sv Acasv-A Acasv-D Acasv-C GASV-A LGASV-C LGASV-D GASV-E Aasv-P Pavev-A Pavsv-B Pavsv-C Aasv-1 Aasv-3 Misv-A Misv-B Misv-F PMSV-4 PMsv-5 PPsv-1 PPsv-2 PPsv-3 PPsv-6 RTSV-2 RTSV-3 PPsv-4 PPSV-5 RTSV-1 NO:21 NO:23 NO:29 NO:31] NO:33] NO:55] NO:25 NO:27 No:35 NO:39 NO:41] NO:43 NO:37 NO:45 NO:49] NO:57] NO:85] NO: 47 NO:51 No:53 NO: 59 NO: 61 NO:63 NO:65 NO: 67 NO: 69 NO:71 NO:73 NO:75 NO:79 NO:81 NO: 77 NO:83 A A A A H H В AA 日日 Н A 88 Н H H H H A В A A Ü A A A B В SEQ SEO SEQ SEQ SEQ SEQ SEQ SEQ SEO SEO SEO SEO SEQ SEO SEO SEQ

### Figure 2 continued

计外角 计电子分类计划计划计划 不知 化苯甲酚 计数字计划计划计划计划计划计划计划 计计计划计划

\*\*\*\*\*\*

ACTGTCAGCAATGATTCCAGCATCCAAGGCAACTGTTTCATCTACCATGTCAAGTTCTCT ACTGTCAGCAATGATTCCAGCATCCAAGGCAACTGTTTCATCTACCATGTCAAGTTCTCT ACTGTCAGCAATGATTCCAGGCATCCAAGGCAACTGTTTCACCTACCATGTCAAGTTCTCT ACTGTCAGCAATGACTCCAGCATCCAAGGCAACTGTTTCACCTACCATGTCAAGTTCTCT ACTGTCAGCAATGATTCCAGCATCCAAGGTAACTGTTTCATCTACAATGTCAAGTTCTCT ACTGTCAGCAATGATTCCAGCATCCAAGGCAACTGTTTCATCTACCATGTCAAGTTCTCT ACTGTCAGCAATGATTCCAGCATCCAAGGCAACTGTTTCATCTACCATGTCAAGTTCTCT ACTGTCAGCAATGATTCCAGCATCCAAGGTAACTGTTTCATCTACCATGTCAAGTTCTCT ACTGTCAGCAATGATTCCAGCATCCAAGGCAACTGTTTCACCTACCATGTCAAGTTCTCT ACTGTCAGCAATGATTCCAGCATCCAAGGCAACTGTTTCACCTACCATGTCAAGTTCTCT ACTGTCAGCAATGATTCCAGCATCCAAGGTAACTGTTTCATCTACAATGTCAAGTTCTCT actgtcagcaatgattccagcatccaaggtaactgtttcatctacaatgtcaagftctct ACTGTCAGCAATGCATCCAAGGCAACTGTTTCACCTACCACGTCAAGTTCTCT ACTGTCAGCAATGATTCCAGCATCCAAGGTAACTGTTTCATCTACAATGTCAAGTTCTCT ACTGTCAGCAATGATTCCAGCATCCAAGGTAACTGTTTCATCTACAATGTCAAGTTCTCT ACTGTCAGCAATGATTCCAGCATCCAAGGCAACTGTTTCATCTACCATGTCAAGTTCTCT ACTGTCAGCAATGATTCCAGGCATCCAAGGCAACTGTTTCATCTACCATGTCAAGTTCTCT ACTGTCAGCAATGATTCCAGCATCCAAGGCAACTGTTTCATCTACCATGTCAAGTTCTCT ACTGTCAGCAATGATTCCAGCATCCAAGGTAACTGTTTCATCTACAATGTCAAGTTCTCT ACTGTCAGCAATGATTCCAGCATGCAAGGCAACTGTTTCATCTACAATGTCAAGTTCTCT ACTGTCAGCAATGATTCCAGCATCCAAGGCAACTGTTTCATCTACAATGTCAAGFTCTCT ACTGTCAGCAATGATTCCAGCATCCAAGGCAACTGTTTCATCTACAATGTCAAGTTCTCT ACTGTCAGCAATGATTCCAGCATCCAAGGCAACTGTTTCACCTACCATGTCAAGTTCTCT ACTGTCAGCAATGATTCCAGCATCCAAGGCAACTGTTTCACCTACCATGTCAAGTTCTCT ACTGTCAGCAATGATTCCAGCATCCAAGGCAACTGTTTCACCTACCATGTCAAGTTCTCT ACTGTCAGCAATGATTCCAGCATCCAAGGCAACTGTTTCACCTACCATGTCAAGTTCTCT actgtcagcaatgattccagcatccaaggtaactgtttcatctacaatgtcaagftctct ACTGTCAGCAATGATTCCAGCATCCAAGGTAGCTGTTTCATCTACAATGTCAAGTTCTCT ACTGTCAGCAATGATTCCAGCATCCAAGGTAACTGTTTCATCTACAATGTCAAGTTCTCT ACTGTCAGCAATGATTCCAGCATCCAAGGCAACTGTTTCACCTACCATGTCAAGTTCTCT ACTGTCAGCAATGATTCCAGCATCCAAGGCAACTGTTTCATCTACCATGTCAAGTTCTCT actgtcagcaatgattccagcatccaaggtaactgtttcatctacaatgtcaagttctct ACTGTCAGCAATGATTCCAGCATCCAAGGTAACTGTTTCATCTACAATGTCAAGTTCTCT ACTGTCAGCAATGATTCCAGCATCCAAGGTAACTGTTTCATCTACAATGTCAAGTTCTCT Ce61-5sv-rep Ce61-7sv-rep PM1Asv-rep PM1Csv-rep GPd58-2sv Ce61-4sv Ce61-3sv LGASV-A Acasv-D LGASV-C Acasv-A Acasv-C LGASV-D LGASV-E Aasv-1 Aasv-3 Aasv-P Pavsv-A Pavsv-B Paver-C Misv-A Misv-B Misv-F PMsv-4 PMBV-5 9-ASG PPsv-1 PPsv-2 PPsv-4 PPSV-5 RTsv-2 Ppsv-3 RTsv-1 NO:23 NO:37] NO:25 NO:27 NO: 29 NO:31 NO:33 NO:35 NO:39 NO:41] No:45 NO:43 NO:47 No:49 NO:51 NO:65] NO:53 NO:55 NO:57] NO:73] NO:59 NO: 61 NO: 63 NO: 69 NO:71 NO:75] NO:67 NO: 79 NO:81 NO: 85 NO:77 NO:83 B 8888 A H A H A A H A A H Н A H Ü H H H A H A A A A H H A B Н SEO SEQ SEQ SEQ SEQ (SEQ SEQ (SEQ SEQ SEO SEQ SEQ SEQ SEQ SEQ SEQ (SEQ SEO SEQ SEQ SEQ SEO SEO SEO SEQ SEO

### Figure 2 continued

\*\*\*\*\*\*\*\*\*\*

Figure 2 continued

使好大块 计电子文字记录文 水水水水水水水水水水水 医苯基水水 化苯基苯法 医苯基苯甲苯苯苯苯苯苯苯苯苯苯

**AACACTGAGCGTCTCTTAGCACGAGATGGAATGCTGATAGGAAACAACTTTATGGCTCTG** AACACTGAGGGTCTCTGCACGÁGATGGAATGCTGATAGGAAACAACTTTTATGGCTCTG CACTCTGAGCGTCTCTTTGCACGAGATGGAATGCTGATAGGAAACAACTTTATGGCTCTG aacactgagcgtcttttgcacgagatggaatgctgataggaaacaactttatggctctg CACTCTGAGCGTCTTTGCACGGGGTGGAATGCTGATAGGAAACAACTTTATGGCTCTG CACTCTGAGCGTCTTTTGCACGGGTGGAATGCTGATAGGAAACAACTTTATGGCTCCG **AACACTGAGCGTCTCTGCACGAGATGGAATGCTGATAGGAAACAACTTTATGGCTCTG** CACTCTGAGCGTCTCTTTGCACGGGGTGGAATGCTGATAGGAAACAACTTTATGGCTCTG CACTCTGAGCGTCTTTTGCACGGGGTGGAATGCTGATAGGAAACAACTTTATGGCTCTG **AACACTGAGCGTCTTTGCACGAGATGGAATGCTGATAGGAAACAACTTTATGGCTCTG** CACTCTGAGCGTCTTTGCACGGGTGGAATGCTGATAGGAAACAACTTTATGGCTCTG AACACTGAGCGTCTTTGCACGAGATGGAATGCTGATAGGAAACAACTTTATGGCTCTG **AACACTGAGCGTCTCTTTGCACGAGATGGAATGCTGATAGGAAACAACTTTTATGGCTCTG AACACTGAGCGTCTTTGCACGAGATGGAATGCTGATAGGAAACAACTTTATGGCTCTG AACACTGAGCGTCTTTGCACGAGATGGAATGCTGATAGGAAACAACTTTATGGCTCTG AACACTGGGCGTCTCTTTGCACGAGATGGAATGCTGATAGGAAACAACTTTATGGCTCTG AACACTGAGCGTCTTTGCACGAGATGGAATGCTGATAGGAAACAACTTTATGGCTCTG AACACTGGGCGTCTTTGCACGAGATGGAATGCTGATAGGAAACAACTTTATGGCTCTG** aacactgagcga ctttatgcacgagatggaatgctgataggaaacaactttatggctctg **AACACTGAGCGTCTTTATGCACGGGATGGAATGCTGATAGGAAACAACTTTATGGCTCTG** aacactgagcgtcttttgcacgagatggagtgctgataggaaacaactttatggctctg CACTCTGAGCGTCTCTTTGCACGGGGTGGAATGCTGATAGGAAACAACTTTTATGGCTCTG CACTCTGAGCGTCTCTTTGCACGGGGTGGAATGCTGATAGGAAACAACTTTAATGGCTCTG CACTCTGAGCGTCTCTTTGCACGGGGTGGAATGCTGATAGGAAACAACTTTATGGCTCTG aacactgagcgtctctttgcacgagatggaatgctgataggaaacaactttatggctctg **AACACTGAGCGTCTCTTTGCACGAGATGGAATGCTGATAGGAAACAACTTTATGGCTCTG AACACTGAGCGTCTCTTTGCACGAGATGGAATGCTGATAGGAAACAACTTTTATGGCTCTG** CACTCTGAGCGTCTCTTTGCACGGGTGGAATGCTGATAGGAAACAACTTTATGGCTCCG **AACACTGAGCGTCTCTTTGCACGAGATGGAATGCTGATAGGAAACAACTTTTATGGCTCTG** aacactgagcgtcttttgcacgagatggaatgctgataggaaacaactttatggctctg CACTCTGAGCGTCTCTTTGCACGGGGTGGAATGCTGATAGGAAACAACTTTATGGCTCTG aacactgagcgtctctctgcacgagatggaatgctaataggaaacaactttatggctctg aacactgagcgtctctttgcacgagatggaatgctgataggaaacaactttatggctctg aacactgagcgtcttttgcacgagatggaatgctgataggaaacaactttatggctctg Ce61-5sv-rep Ce61-7sv-rep PM1Asv-rep PM1Csv-rep NO: Pavsv-C NO: Pavsv-B 3Pd58-2sv Ce61-4sv Ce61-3sv Acasv-C Acasv-D LGASV-A Acasv-A LGASV-C LGASV-D GABV-E A-VSVe Aasv-P Misv-A NO:19]: Aasv-1 Aasv-3 Misv-B Misv-F PMsv-5 PMsv-4 PPsv-1 PEV-2 PPsv-3 PPSV-4 PPSV-5 PPsv-6 RTsv-1 RTsv-2 RTSV-3 NO:23] NO:25] NO:29 NO:31 NO:35 NO:39 NO:41] NO:49] NO:27 NO:37 NO:43 NO:45] NO:47] NO:51 NO:53 NO:65] No:69] NO:33 NO:55 NO:57 NO:59 NO:61 NO: 63 NO: 67 NO:71 NO:73 NO:75] NO:79 NO:83 NO:81 NO:77 A A A A A A A A B Н H A 自自自 A 品品 A H A H A H H 日日 H B H SEO SEQ SEQ SEQ. CES] SEQ SEQ SEQ SEO SEO SEQ SEQ SEO SEQ SEQ

### Figure 2 continued

化妆妆妆妆妆妆妆妆妆妆妆妆妆妆妆妆妆妆妆妆妆妆妆 化水妆妆 女 大大大大大

TTGGAAGGAGGTGGTCATTATTTGTGTGAATTCAAATCTACTTACAAGGCAAAGAAG TTAGAAGGAGGTGGTCACTATTTGTGTGAATTCAAATCTACTTACAAGGCAAGGAAG TTAGAAGGAGGTGGTCACTATTTGTGTGAATTCAAATCTACTTACAAGGCAAAGAAG TTAGAAGGAGGGGTCACTATTTGTGTGGATTCAAAACTACTTACAAGGCAAAGAAG TTAGAAGGAGGTGACTATTTGTGTGAATTCAAATCTACTTACAAGGCAAGGAAG TTAGAAGGAGGGGTCACTATTTGTGTGAATTCAAAACTACTTACAAGGCAAAGAAG TTAGAAGGAGGTGGTCACTATTTGTGTGAATTCAAATCTACTTACAAGGCAAGGAAG TINGAAGGAGGNGGTCANTATITGTGTGAATTCAAATCTACTTACAAGGCAAAGAAG ITAGAAGGGGGGGTCACTATTTGTGTGAATTCAAAACTACTTACAAGGCAAAGAAG TTAGAAGGAGGGGTCACTATTTGTGTGAATTCAAAACTACTTACAAGGCAAAGAAG ITGGAAGGAGGTGGTCATTATTTGTGTGAATTCAAATCTACTTACAAGGCAAAGAAG TIGGAAGGAGGIGGICAITATITIGIGTGAATICAAATCTACTIACAAGGCAAAGAAG ITAGAAGGAGGGGTCACTATTTGTGTGAATTCAAAACTACTTACAAGGCAAAGAAG TTGGAAGGAGGTGGTCATTATTTGTGTGAATTCAAATCTACTTACAAGGCAAAGAAG ITGGAAGGAGGTGGTCATTATTTGTGTGAATTCAAATCTACTTACAAGGCAAAGAAG ITAGAAGGAGGTGGTCACTATTTATGTGAATTCAAATCTACTTACAAGGCAAGGAAG TAGAAGGAGGTGGTCACTATTTGTGTGAATTCAAATCTACTACAAGGCAAGGAAG TAGAAGGAGGTGGCCACTATTTGTGTGAATTCAAATCTACTTACAAGGCAAGGAAG FTGGAAGGAGGTGGTCATTATTTGTGTGAATTCAAATCTACTTACAAGGCAAAGAAG ITGGAAGGAAGTGGTCATTATACCTGTGAATTCAAATCTACTTACAAGGCAAAGAAG ltggaaggagg-----Gcaaagaag TTGGAAGGAGGTGGTCATTATTTGTGTGAATTCAAATCTACTTACAAGGCAAAGAAG TTAGAAGGAGGCGGTCACTATTTGTGTGGATTCAAAACTACTTACAAGGCAAAGAAG

TAGAAGGAGGCGGTCACTATTTGTGTGGATTCAAAACTACTTACAAGGCAAAGAAG ttagaaggaggcggtcactatttgtgtggattcaaaactacttacaaggcaaagaag

TAGAAGGAGGCGGTCACTATTTGTGTGGATTCAAAACTACTTACAAGGCAAAGAAG TGGAAGGAGGTGGTCATTATTTGTGTGAATTCAAATCTACTACAAGGCAAAGAAG TGGAAGGAGGTGGTCATTATTTGTGTGAATTCAAATCTACTACAAGGCAAAGAAG TGGAAGGAGGTGGTCATTATTTGTGTGAATTCAAATCTACTACAAGGCAAAGAAG TAGAAGGAGGGGTCACTATTTGTGTGAATTCAAAACTACTTACAAGGCAAAGAAG TAGAAGGAGGTGGTCACTATTTGTGTGAATTCAAATCTACTTACAAGGCAAGGAAG TGGAAGGAGGTGGTCATTATTTGTGTGAATTCAAATCTACTTACAAGGCAAAGAAG TGGAAGGAGGTGGTCATTATTTGTGTGAATTCAAATCTACTTACAAGGCAAAGAAG TGGAAGGAGGTGGTCATTATTTGTGTGAATTCAAATCTACTTACAAGGCAAAGAAG

SEQ ID NO:19] Aasv SEQ ID NO:21] Aasv SEQ ID NO:23] Aasv SEQ ID NO:23] Acas SEQ ID NO:23] Acas SEQ ID NO:23] Acas SEQ ID NO:29] Acas SEQ ID NO:33] Cecl. SEQ ID NO:33] Cecl. SEQ ID NO:33] Cecl. SEQ ID NO:33] GPAGS SEQ ID NO:33] GPAGS SEQ ID NO:45] LGAS SEQ ID NO:45] LGAS SEQ ID NO:45] LGAS SEQ ID NO:45] LGAS SEQ ID NO:45] MASV SEQ ID NO:53] MASV SEQ ID NO:53] MASV SEQ ID NO:53] PRICE SEQ ID NO:53] PRICE SEQ ID NO:53] PRICE SEQ ID NO:53] PRICE SEQ ID NO:63] RICE PUBLIC SEQ ID NO:63] RICE PUBLIC SEQ ID NO:63] RICE PUBLIC PUBLI	4	PAG	AAG	AAG	AAG	AAG	AAG	AAG	AAG	AAG	AAG	AAG	AAG.	AAG	AAG	AAG	AAGI	PAG	AAGI	AAG	AAGI	AAGI	AAGI	AAGI	AAGI	AAGI	AAGI	AAGI	AAGT	AAGT	AAGT	AAGT	AAGT	AAGT
	EQ ID NO:19]	ID NO:21] Aasv-	ID NO:23] Aasv-	ID NO:25] Acasv-	ID NO:27] Acasv-	ID NO:29] Acasv	ID NO:31]	ID NO:33] Ce61-4	ID NO:35] Ce61-5sv	ID NO:37] Ce61-7sv	ID NO:39] GPd58-2sv	ID NO:41] LGASV-	ID NO:43]	ID NO:45] LGASV	ID NO:47] LGASV-	ID NO:49] Misv-	ID NO:51]	ID NO:53]	ID NO:55] PMIASV-	ID NO:57] PMICSV-	ID NO:59] PMBV-4	ID NO:61] PMSV-	ID NO:63]	ID NO:65]	ID NO:67]	ID NO:69] PPSV-	ID NO:71] PPSV-	ID NO:73 PPsv-	ID NO:75] Pavsv-	ID NO:77] Pavsv-	ID NO:79] Pavsv-	ID NO:81] RISV	EQ ID NO:83] RISV-	[SEQ ID NO:85] RTSV-3

igure 2 continued

不不不不不 水水不不不

ATGTTGACCGCAAACTGGATGTAACCAATCACAAC ATGTTGACCGCAAACTGGATGTAACCAATCACAAC GTGTTGACCGCAAACTGGATGTAACCAATCACAAC ATGTTGACCGCAAACTGGATGTAACCAATCACAAC

ATGTTGACCGCAAACTGGATGTAACCAATCACAAC

**ATGTTGACCGCAAGTTGGATGTAACCAATCACAAC** 

ATGTTTACAGCACCATTCATGTAACCAATCACAAC ATGTTGACCGCAAACTGGATGTAACCAATCACAAC ATGTTGACCGCAAATTGGATGTAACCAATCACAAC ATGTTGACCGCAAATTGGATGTAACCAATCACAAC

ATGTTGACCGCAAACTGGATGTAACCAATCACAAC
ATGTTGACCGCAAATTGGATGTAACCAATCACAAC
ATGTTGACCGCAAATTGGATGTAACCAATCACAAC
ATGTTGACCGCAAATTGGATGTAACCAATCACAAC
ATGTTGACCGCAAACTGGATGTAACCAATCACAAC
ATGTTGACCGCAAACTGGATGTAACCAATCACAAC

ITGTTGACCGCAAGTTGGATGTAACCAATCACAAC **ATGTTGACCGCAAATTGGATGTAACCAATCACAAC NTGTTGACCGCGAATTGGATGTAACCAATCACAAC ITGTTGACCGCAAATTGGATGTAACCAATCACAAC ATGTTGACCGCAAACTGGATGTAACCAATCACAAC NIGITGACCGCAAACTGGATGTAACCAATCACAAC ITGTTGACCGCAAACTGGATGTAACCAATCACAAC** ITGITGACCGCAAACTGGATGTAACCAATCACAAC ITGTTGACCGCAAATTGGATGTAACCAATCACAAC ITGITGACCGCAAATTGGATGTAACCAATCACAAC IGTTGACCGCAAATTGGATGTAACCAATCACAAC ITGTTGACCGCAAACTGGATGTAACCAATCACAAC ITGITIGACCGCAAACTGGATGTAACCAATCACAAC TGTTGACCGCAAATTGGATGTAACCAATCACAAC TGTTGACCGCAAATTGGATGTAACCAATCACAAC ITGTTGACCGCAAATTGGATGTAACCAATCACAAC

atgitgaccgcaaactggatgtaaccaatcacaac atgitgaccgcaaattggatgtaaccaatcacaac

(SEQ ID NO:19] A	Aasv-1	CCTGTGATGATGCCAGGGTATCACTZ
[SEQ ID NO:21] A	Aasv-3	CCTGTGAAGATGCCAGGGTATCACTZ
[SEQ ID NO:23] A	Aasv-P	CCTGTGAGGATGCCAGGGTATCACT
[SEQ ID NO:25] A	Acasv-A	CCTGTGAAGATGCCAGGGTATCATT
[SEQ ID NO:27] A	Acasv-C	CCTGTGAAGATGCCAGGGTATCACTC
[SEQ ID NO:29] A	Acasv-D	CCTGTGAAGATGCCGGGGTATCATT
ID NO:31]	Ce61-3sv	CCTGTGAAGATGCCAGGGTATCACTA
[SEQ ID NO:33] C	Ce61-4sv	CCTGTGATGATGCCAGGGTATCACTA
ID NO:35]	Ce61-5sv-rep	CCTGTGAAGATGCCAGGGTATCATT
ID NO:37]	Ce61-7sv-rep	CCTGTGAAGATGCCAGGGTATCATT
ID NO:39]	GPd58-2sv	CCTGTGATGATGCCAGGGTATCACTA
ID NO:41]	LGASV-A	CCTGTGATGATGCCAGGGTATCACTA
ID NO:43]	LGAsv-C	CCTGTGAAGATGCCAGGATATCATTA
[SEQ ID NO:45] L	LGASV~D	CCTGTGATGATGCCAGGGTATCACTA
ID NO:47]	LGASV-E	CCTGTGATGATGCCAGGGTATCACTA
ID NO:49]	Misv-A	CCTGTGAAGATGCCAGGGTATCACTA
[SEQ ID NO:51] M	Misv-B	CCTGTGAAGATGCCAGGGTATCACTA
ID NO:53]	Misv-F	CCTGTGAAGATGCCAGGGTATCACTA
REQ ID NO:55]	PM1Asv-rep	CCTGTGATGATGCCAGGGTATCACTA
ID NO:57]	PM1Csv-rep	CCTGTGATGATGCCTGGATATCACTA
ID NO:59]	PMSV-4	CCTGTGATGATGCCAGGGTATCACTA
ID NO:61]	PMsv-5	CCTGTGATGATGCCAGGGTATCACTA
ID NO:63]	PPsv-1	CCTGTGAAGATGCCAGGGTATCATTA
ID NO:65]	PPsv-2	CCTGTGAAGATGCCAGGGTATCATTA
ID NO:67]	PPsv-3	CCTGTGAAGATGCCAGGGTATCATTA
ID NO:69]	PPsv-4	CCTGTGAAGATGCCAGGGTATCATTA
ID NO:71]	PPsv-5	CCTGTGATGATGCCAGGGTATCACTA
	Psv-6	CCTGTGATGATGCCAGGGTATCACTA
ID NO:75]	Pavsv-A	CCTGTGATGATGCCAGGGTATCACTA
ID NO:77]	Pavsv-B	CCTGTGAAGATGCCGGGGTATCATTA
[67:0N CI	Pavsv-C	CCTGTGAAGATGCCAGGGTATCACTA
ID NO:81]	RTsv-1	CCTGTGATGATGCCAGGGTATCACTA
[SEQ ID NO:83] R	RTsv-2	CCTGTGATGATGCCAGGGTATCACTA
[SEQ ID NO:85] R	RTsv-3	CCTGTGATGATGCCAGGGTATCACTA

### Figure 2 continued

\*\*\*\*\*\*\*\*\*\*\*\*

水水 水水水水 水 水水水水水 水水 化水水水水 化水水水水水

15/94

	NO:19] NO:21]	Aasv-1 Aasv-3	AAGGATTACACTTCCGTTGAGCAGTGTGAAALTTTCCATTGCACGCAAACCTGTGGTCGCC AAGGATTACACTTCCGTTGAAGCAGCGTGAAATTTCCATTGCACGCAAACCTGTGGTCGCC
	No:23]	Aasv-P	AAGGATTACACTTCCGTTGAGCAGTGTGAAATTTCCATTGCACGCAAACCTGTGGTCGCC
(SEQ ID	NO:25]	Acasv-A	AAGGATTACATTTCCGTTGAGCAGTGTGAAATTTCCATTGCACGCAAACCTGTGGTCGCC
(SEQ ID	NO:27]	Acasv-C	AAGGATTACACTTCCGTTGAGCAGCGTGAAATTTCCATTGCACGCAAACCTGTGGTCGCC
CI OES]	NO:29]	Acasv-D	AAGGATTACACTTCCGTTGAGCAGTGTGAAATCTCCATTGCACGCAAACCTGTGGTCGCC
(SEQ ID	NO:31]	Ce61-3sv	AAGGATTACACTTCCGTTGAGCAGCGTGAAATTTCCATTGCACGCAAACCTGTGGTCGCC
•	NO:33]	Ce61-4sv	AAGGATTACACTTCCGTTGAGCAGTGTGAAATATCCATTGCACGCAAACCTGTGGTCGCC
	NO:35]	Ce61-5sv-rep	AAGGATTACACTTCCGTTGAGCAGTGTGAAATTTCCNNTNCACGCAAACCTGTGGGTCGCC
(SEQ ID	NO:37]	Ce61-7sv-rep	AAGGATTACACTTCCGTTGAGCAGTGTGAAATTTCCATTGCACGCAAACCTGTGGTCGCC
(SEQ ID	NO:39]	GPd58-2sv	AAGGATTACACTTCCGTTGAGCAGTGTGAAATTTCCATTGCACGCAAACCTGTGGTCGCC
	NO:41]	LGASV-A	AAGGATTACACTTCCGTTGAGCAGTGTGAGAATTTTCCATTGCACGCAAACCTGTGGTCGCC
	NO:43]	LGASV-C	AAGGATTACACTTCCGTTGAGCAGTGTGAAATTTCCATTGCACGCAAACCTGTGGTCGCC
	NO:45]	LGASV-D	AAGGATTACACTTCCGTTGAGCAGTGTGAAATTTTCCATTGCACGCAAACCTGTGGTCGCC
	NO:47]	LGASV-E	AAGGATTACACTTCCGTTGAGCAGTGTGAAATTTCCATTGCACGCAAACCTGTGGTCGCC
	NO:49]	Misv-A	AAGGATTACACTTCCGTTGAGCAGCGTGAAATTTCCATTGCACGCAAACCTGTGGTCGCC
H	NO:51]	Misv-B	AAGGATTACACTTCCGTTGAGCAGCGTGAAATTTCCATTGCACGCAAACCTGTGGTCGCC
A	NO:53]	Misv-F	AAGGATTACACTTCCGTTGAGCAGCGTGAAATTTCCATTGCACGCAAACCTGTGGTCGCC
A	NO:55]	PM1Asv-rep	AAGGATTACACTTCCGTTGAGCAGTGTGAAATTTCCATTGCACGCAAACCTGTGGTCGCC
A	NO:57]	PM1Csv-rep	AAGGAITACACTICCGITGAGCAGTGTGAAATTICCAITGCACGCAAACCTGTGGTCGCC
A	NO:59]	PMsv-4	AAGGATTACACTTCCGTTGAGCAGTGTGAGATTTCCATCGCACGCA
A	NO:61]	PMsv-5	AAGGATTACACTTCCGTTGAGCAGTGTGAGATTTCCATCGCACGCA
H	NO:63]	PPsv-1	AAGGATTACATTTCCGTTGAGCAGTGTGAAATTTCCATTGCACGCAAACCTGTGGTCGCC
A	NO:65]	PPsv-2	AAGGATTACATTTCCGTTGAGCAGTGTGAAATTTCCATTGCACGCAAACCTGTGGTCGCC
A	NO:67]	PPsv-3	AAGGATTACATTTCCGTTGAGCAGTGTGAAACTTTCCATTGCACGCAAACCTGTGGTCGCC
П	NO:69]	PPsv-4	AAGGATTACATTTCCGTTGAGCAGTGTGAAATTTCCATTGCACGCAAACCTGTGGTCGCC
A	NO:71]	PPsv-5	AAGGATTACACTTCCGTTGAGCAGTGTGAAATTTTCCATTGCACGCAAACCTGTGGTCGCC
A	NO:73]	PPsv-6	AAGGATTACACTTCCGTTGAGCAGTGTGGAATTTCCATTGCACGCAAACCTGTGGGTCGCC
A	NO:75]	Pavsv-A	AAGGATTACACTTCCGTTGAGCAGTGTGAAATTTCCATTGCACGCAAACCTGTGGTCGCC
A	NO:77]	Pavsv-B	AAGGATTACACTTCCGTTGAGCAGTGTGAAATCTCCATTGCACGCAAACCTGTGGTCGCC
H	NO:79]	Pavsv-C	AAGGATTACACTTCCGTTGAGCAGCGTGAAATTTTCCATTGCACGCAAACCTGTGGTCGCC
A	NO:81]	RTsv-1	AAGGATTACACTTCCGTTGAGCAGTGTGAAATTCCCATTGCACGCAAACCTGTGGTCGCC
A	NO:83]	RTsv-2	AAGGATTACACTTCCGTTGAGCAGTGTGAAATTTCCATTGCACGCAAACCTGTGGTCGCC
[SEQ ID	NO:85]	RTsv-3	AAGGATTACACTTCCGTTGAGCAGTGTGAAATTTCCATTGCACGCAAACCTGTGGTCGCC

continued
2 co
Figure

Aasv-1	
Aasv-3	
Aasv-P	TGA
Acasv-A	1
Acasv-C	
Acasv-D	
Ce61-3sv	
Ce61-4sv	*****************
Ce61-5sv-rep	
Ce61-7sv-rep	
GPd58-2sv	
LGASV-A	
LGASV-C	
LGASV-D	
LGASV-E	1
Misv-A	
Misv-B	
Misv-F	
PM1Asv-rep	
PM1Csv-rep	
PMsv-4	TGACGTTTTTCAGAGTCAAATCAAGGCACAA
PMsv-5	TGACGTTTTTCAGAGTCAAATCAAGGCACAA
PPsv-1	
PPsv-2	
PPsv-3	
PPsv-4	
PPsv-5	
PPsv-6	8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8
Pavsv-A	7 3 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7
Pavsv-B	
Pavsv-C	**********************
RTsv-1	
RTsv-2	
RTsv-3	

### Figure 3

(SEC ID NO.88)	A C - S H E A	
	Aams-2.pcp	ATTGLEVIOLISTS FLEGAL*ELFANKCSALLINMSVIATOMTYKVYMSGTVNGHYFEVE ATTGLEVANTSTSEPPERATEPANDGSITT TIMESTFANDAMASSTSEP
[SEQ ID NO:92]	Aams-5.pep	STABLISON TOOM AND THOUGHT CONTINUES OF THE TRANSPORT OF
(SEQ ID NO:94)	Aams-6.pep	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
(SEQ ID NO:96]	Aams-A.pep	SGIATOMIYKUMS@UTUMS@UTUMS@UTUMS@UTUMS@UTUMING#UTUMS@UTUMING#UTUMS@UTUMING#UTUMP@UTUMP###################################
(SEQ ID NO:98]	Aams-B.pep	SVIATOMIYKVYMSGTWNGHYFEVRGDGKGKPVEGEOTTAFLAVTAGEDLESPOO
[SEQ ID NO:100]	Acams-2.pep	ODJECTTORWING PRONT AND A TROOT TO THE TROOT
	Acams-3.pep	3
A	Acams-4.pep	TATTONE TO CONTACT TO THE ACT.
A	Acams-5.pep	47447444666666666666666666666666666666
H	Cems-F.pep	
A	Cems-G.pep	EVER VERNING TO COMPANY AND CO
A	Cems-H.pep	GARRANDUNALDONALNAMOUNTAANA
ü	Cems-I.pep	EAREARINALOS WAADAALWOLD INSTITUTE TO THE TOTAL THE TOTAL TO THE TOTAL THE TOTAL TO THE TOTAL TH
H	LGAms-5.pep	CARTINONALOS MANAMANAMOLIS INTERPORTANTA CONTRACTOR OF THE CONTRAC
[SEG ID NO:118]	LGAms-6.pep	TATTING ADOLES AND THE COMPANY OF TH
(SEQ ID NO:120)	Mi68Dms.pep	1
[SEQ ID NO:122]	Mims-A.pep	1
[SEQ ID NO:124]	Mims-B.pep	
[SEQ ID NO:126]	Mims-C.pep	
[SEQ ID NO:128]	PMms-A.pep	•
(SEO ID NO:130)	PMms-B.nen	HARRIAN AND THE ACTION AND THE ACTION AND ACTION A
A	Diving Con Dead	EASTWSV.LAT.QMLY.KVYMSGTVNGHYPEVE
A	DMms-D. nen	STATE OF THE STATE
E	DMng Troub	!
1	משליבר ביוווים	!
} }	The care the part	!
1 6	rrds7-2ms.pep	
3 6	ged: -/coa/a	1
1. F	reds/-4ms.pep	;
3 5	ded T-smar	1 1 1
9 6	1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 -	1 1 1 1
3 6	190 - 1 - SIII 4 4	SVIATYKVYMSGTVNGHYFEVE
1	ded o-swar	SVIATOMIYKVYMSGTVNGHYFEVE
8	Pav5ms.pep	EAST- SAIDMIXKOAMSCIANGHYPEVE
H	Pavms-2.pep	
H	Pavms-3.pep	;
A	Pavms-4.pep	1 1
	RIms-1.pep	1
H	RTms-2.pep	GAG THIDWAYDONA MALWOL VANAWA THE TARREST TO THE TRANSPORT OF THE TRANSPOR
ü	RIms-5.pep	GARITANA DOMANA MANAMANA LA CARACTERIA DE LA CARACTERIA D
SEQ ID NO:168]	RTms-6.pep	GAG STUDWATESWADXALMOLA IASA COLAR GAG STUDWATES AND STUDW
		₹ ^ 1   1 = 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

		SECTION CONCRETED TO THE TRANSPIRE TO SECTION TO SECTIO						n.	GDGKGKPYEGEQTVKLTVTKGGPLDFAWDILSPOCOYGNIPFTKYPFDVPDVVVPSFF						0.																				Ω					*
Aams-2.pep	Aams-5.pep	Aams-6.pep	Aams-A.pep	Aams-B.pep	Acams-2.pep	Acams-3.pep	Acams-4.pep	Acams-5.pep	Cems-F.pep	Cems-G.pep	Cems-H.pep	Cema-I.pep	LGAms-5.pep	rgyms-6.pep	Mi68Dms.pep	Mims-A.pep	Mims-B.pep	Mims-C.pep	PMms-A.pep	PMms-B.pep	PMms-C.pep	PMms-D.pep	PMms-E.pep	PPd57-1ms.pep	PPd57-2ms.pep	PPd57-3.pep	PPd57-4ms.pep	PPms-1.pep	PPms-2.pep	PPms-E.pep	PPms-G.pep	Pav5ms.pep	Pavms-2.pep	Pavms-3.pep	Pavms-4.pep	RTms-1.pep	RTms-2.pep	RIms-5.pep	KIms-6.pep	* * * * *
(SEQ ID NO:88)		ü	ü	B	ü	ü	ü	H	H	ij	H	H	吕	8	H	H		£.	H	H	H	ü	Ü	H	ü	H		유	ij	H	H	B	A	ដ	H		A	H	[SEQ ID NO:168]	

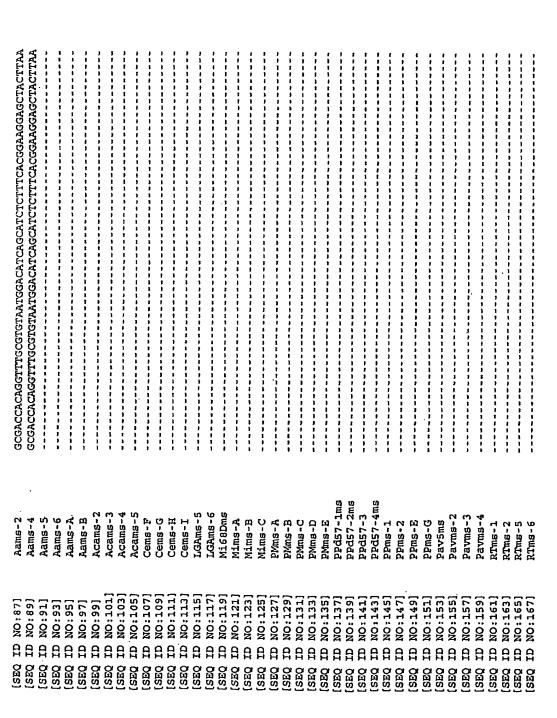
### Figure 3 continued

### 3FTWER IMNFEDGAVCTVSNDSSI QGNCFI YHVKFSGLNF PPNGPVMQKKTQGWEPHS sglnfppngpvmokktogwephserlfardgmlignnfmalkleggghylcefktitvk gytwerimnfedgavctvsndssiqgncfiyhvkfsginfppngpvmokktqgwepnt gytwerimnfedgavctvsndssiqgncfihhvkfsglnfppngpvmqkktqgwepnt GFTWERIMNFEDGAVCTVSNDSSIQGNCFTYHVKFSGLNFPPNGPVMQKKTQGWEPHS GFTWERLIMNFEDGAVCTVSNDSSIQGNCFTYHVKFSGLNFPPNGPVMQKKTQGWEPHS GFTWERIMNFEDGAVCTVSNDSSIQGNCFTYHVKFSGLNFPPNGPVM\*KKTQGWEPHS GFTWDRIMDFEDGAVCTVSNDSSIQGNCFIYHVKFSGLNFPPNGPVMQKKTQGWEPNT GFTWER IMNFENGAVCTVSNDSSIQGNCFTYHVKFSGLNFPPNGPVMQKKTQGWEPHS GFTWERIMNFEDGAVCTVSNGSSIQGNCFTYHVKFSGLNFPPNGPVMQKKTQGWEPHS GFTWERIMNFEDGAVCTVSNDSSIQGNCFTYHVKFSGLNFPPNGPVMQKKTQGWEPHS GYTWERIMNFEDGAVCTVSNDSSIQGNCFIYNVKFSGLNFPPNGPVMQKKTQGWEPNT GFTWERIMNFEDGAVCTVSNDSSIQGNCFI YHVKFSGLMFPPNGPVMQKKTQGWEPHS GFTWERIMNFEDGAVCTVSNDSSIQGNCFIYHVKFSGLNFPPNGPVMQKKTQGWEPHS GFTWER IMNFEDGAVCTVSNDSSI QGNCFTYHVKFSGLNFPPNGPVMQKKTQGWEPHS GFTWDR.IMNFEDGAVCTVSNDSSIQGNCFIYHVKFSGLNFPPNGPVMQKKTQGWEPNT GFTWER IMNFEDGAVCTVSNDSS I QGNCFTYHVKFSGLNFPPNGPVMQKKTQGWEPHS GFTWEGIMNFEDGAVCTVSNDSSIQGNCFTYHVKFSGLNFPPNGPVMQKKTQGWEPHS gytwer imnfedgavctvsndssi ogncfi yhvkfsginfppngpvmokktoggwednt gytwerimnfedgavctvsndssiogncfiyhvkfsglnfppngpvmokktoggwednt **GYTWEGIMNFEDGAVCTVSNDSSIQGNCPIYHVKFSGLNFPPNGPVMQKKTQGWEPNT** gytwerimnfedgavcavsndssiqgncfiyhvkfsglnfpengpvmqkktqgwepnt gftwerinnfedgavctvsndssiqgncftyhvkfsginfppngpvmokktogwephs 3FTWERIMNFEDGAVCTVSNDSSIQGNCFIXHVKFSGLNFPPNGPVMQKKTQGWEPHS 3FTWERIMNFEDGAVCTVSNDSSIQGNCFIYHVKFSGLNFPPNGPVMQKKTQGWEPHS GFTWERIMNFEDGAVCTVSNDSSIQGNCFIYHVKFSGLNFPPNGPVMQKKTQGWEPHS SGLNFPPNGPVMQKKTQGWEPHSERLFARDGMLIGNNFWALKLEGGGHYLCEFKTTTK **GYTWERIMNFEDGAVCTVSNDSSIQGNCFIYHVKFSGLNFPPNGPVMQKKTQGWEPNT** GFTWERIMNFEDGAVCTVSNDSSIQGNCFTYHVKFSGLNFPPNGPVMQKKTQGWEPHS gytwer imnfedgavctvsndssi Qgncf i yhvkfsglnfppngpvmQkktQgwepnt GFTWERIMNFEDGAVCTVSNDSSIQGNCFTYHVKFSGLNFPPNGPVMQKKTQGWEPHS **GYTWERIMNFEDGAVCTVSNDSSIQGNCFIYNVKFSGLNFPPNGPVMQKKTQGWEPNT** gytwer imnfedgavctvsndssiogncfiynvkfsglnfppngpvmokktogwepnt gytwerimkfedgavctvtndssmogncfiynvkfsglnfppngpvnokktoggwepnt gytwer imnfedgavctvsndss i Qgncf i ynvkfsglaf pngpvmokktoggwe pnt gytwer imnfedgavctvsndss i ogncf i ynv kfsglnfp pngpvmokktogwepnt **GYTWERIMNFEDGAVCTVSNDSSIQGNCFIYHVKFSGLNFPPNGPVTQKKTQGWEPNT** gytwer imnfedgavctvsndssiqgncfiyhvkfsginfppngpvmokktogwepnt gytwerimnfedgavctvsndssiogncfiyhvkfsglnfppngpvmokktoggwednt gytwer imkfedgavctvsndssmogncf i ynvkfsginfppngpvmokktogwednt \*\*\*\* 经经济证据 化水水水 PPdS7-4ms.pep PPd57-1ms.pep PPd57-2ms.pep 化放性化 化化 化妆 LGAms-6.pep Mi68Dms.pep PPd57-3.pep Pavme-2.pep Pavms-3.pep LGAms-5.pep Pavms-4.pep Acams-2.pep Acams-5.pep Acams-3.pep Acams-4.pep Cems-G.pep Cems-H.pep PMms-E.pep PPms-2.pep PPms-E.pep Cems-F.pep Cems-I.pep Mims-B.pep Mims-C.pep PMms-A.pep PMms-B.pep PMms-C.pep PMms-D.pep PPms-1.pep Pms-G.pep RIms-1.pep RIms-5.pep lams-4.pep Aams-5.pep dad.9-smet Aams-A.pep Aams-B.pep Mims-A.pep Pav5ms.pep RIms-2.pep RIms-6.pep NO:152] NO:114 NO:116 NO:118 NO:120 NO:122] NO:124] NO:126] NO:128] NO:130 NO:138] NO:140] NO:142] NO:144] NO:146] NO:148 NO:150 NO:158 NO:160 NO:166 NO:168 NO: 100 NO:102 NO: 104 NO:106 NO:108 NO:110 NO:112 NO:132 NO:134 NO:136 NO:154 NO:156 No:162 NO:164 NO:90 NO:96] NO: 92 NO:94] NO: 98 品品 A B 888 A 8888 H 88 ü 88 88 H 888 Ü ü ដ H A a H A В 88 H Casion of San Ca SEO

<b>&gt;</b> :	> 5	> 5	> 1			. 5							. ~																										
ERLFARDGMLIGNNEMALKLEGGGHYLCEFKTTYKAKKPVKMPGYHYVDRKLDV	ERLEARDGMI. IGNNEMALKI. RGGGHY. CREKTTY KAKKEV KMPGYHYNDRKLDV ERLEARDGMI. IGNNEMAI. KI. RGGGHYI. CREKTTY VA KY DIXMDGYHYNDBYT. DIX	ERLFARDGMLIGNNFMALKLEGGGHYL, CEFKTTYKAKKKDVKMPCYHYMBKLDN	AKKPVKMPGYHYVDRKIDVINHNKDYTSVEOCEISTABKPVVA	AKKPVKMPGYHYVDRKCDVINHNKDYTSVEOCEISIARNPVVA	ERLFARDGMLIGNNFMALKLEGGGHYLCEFKSTYKA KR PVKMPGYHVVDBKT DV	ERLFARDGMLIGNNFMALKLEGGGHYI.CBFKSTYKA KKPVKMPGYHYVIDV	ERLFARDGMLIGNNEMALKIEGGGHYI/CEFKSTYKAKKPVKMPGYHYNDBKT.DV		ERL PARGGMLIGNN FMALKLEGGGHYLCGFKTTYKAKKPVKMPGYHYVDRKTDV	erllargomilgnnfmalkleggghylcgpkttykakkpykmpgyhyndrytnv	ERLFARGOMLIGNNEMALKLEGGGHYLCGFKTTYKAKKI.VKMPGYHYVDRKT.DV	ERLFARGGMLI GNNFMALKLEGGGHYLCEFKTTYKAKKPVKOPGYHYADRKT.DV	ERLLARDGMLI GNNEMALKLEGGGHYLCEFKSTYKAKKPVKMPGYHYVDRKT.DV	ERLFARDGMLLGNNFMALKLEGGGHYLCEFKTTYKAKKPVKMPGYHYVDRKT.DV	ERLFARGGMLIGNNFMALKLEGGGHYLCEFKTTYKAXKPVKMPGYHYVDRKTDV	erlfarggmlignnfmalkleggghylcefkttykakkpykmpgyhyvdrktdv	ERLFARGGMLI GNNFMALKLGGGGHYLCEFKTTYRAKKPVKMPGYHYVDBKT.DV	ERL FARGGMLI GNNFMALKLEGGGHYLCEFKTTYKAKKPVKMPGYHYYDBYT.DV	ERLYARDGMLIGNNEMALKLEGGGHYLCEFKSTYKAKKPVMMPGYHYVNBKT.)	ERL PARDGMLIGNNFMALKLEGGGHYLCEFKSTYKAKKPVMMPGYHYVDRYT.DV	GRLYARDGMLIGNNFMALKLEGGGHYTCEFKSTYKAKKBUMMBGYHYYDBKTDW	erl fardgmlignnemalkseggghylcefkstykakkpymmpgyhyvdrktidv	ERLFARDGMLIGNNEMALKLEGGGHYLCEFKSTYKAKKPVMMPGYHYVDRKIDV	erlyardgmlignnfmalkdeggghytcefkstykakkrynmpgyhyvdrkidv	erlfardgalignnfmalkleggghylcefkittykakkpykapgyhyddrhod	ERLFARDGMLI GNNFMALKLEGGGHYLCEFKTTYKAKKPVKMPGYHYVDRKLDV	ERLFARDGMLI GNNFMALKLEGGGHYLCEFKSTYKAKKPVMMPGYHYVDRKT,DV	ERLFARGGMLI GNNFMALKLEGGGHYLCGFKTTYKAKKPVKMPGYHYVDRKLDV	erllardgmliginnemalkleggghylcefkttykakkevkmpgyhyvdrkldv	ERLFARGGMLIGNNFMALKLEGGGHYLCGFKTTYKAKKPVKMPGYHYVDRKLDV	erl fargomlignnemalkleggghylcgfkttykakkpykmpgyhyvdrktdv	ERLFARDGMLI GNNFMALKLEGGGHYLCEFKSTYKAKKTVKMPGYHYVDRKT.DV	ERLFARDGLLIGNNEMALKLEEGGHYLCEFKSTYKAKKTAKMPGYHYVDRKT.DV	ERLFARDGMLIGNNFMALKLEGGGHYLCEFKSTYKAKKTVKMPGYHYVDRKT,DV	erlfardgmlignnfmalkleggghylcefkstykakktvkmpgyhyydrklyv	ERLFARDGMLIGNNFMALKLEGGGHYLCEFKSTYKAKKPVKMPGYHYVNRKT.NV	ERLFARDGMLIGNNFMALKLEGGGHYLCEFKSTYKAKKPVKMPGYHYVDRKLDV	ERLFARGGML I GNNFMALKLEGGGHYL CEFKTTYKAKKPVKMPGYHYVDRKCDV	ERLFARDGMLIGNNFMALKLEGGGHYLCEFKSTYKAKKPVKMPGYHYVDRKLDV
Aams-2.pep Aams-4.pep	Aams-5.pep	Aams-6.pep	Aams-A.pep	Aams-B.pep	Acams-2.pep	Acams-3.pep	Acams-4.pep	Acams-5.pep	Cems-F.pep	Cems-G.pep	Cems-H.pep	Cems-I.pep	LGAms-5.pep	LGAms-6.pep	Mi68Dms.pep	Mims-A.pep	Mims-B.pep	Mims-C.pep	PMms-A.pep	PMms-B.pep	PMms-C.pep	PMms-D.pep	PMms-E.pep	PPd57-1ms.pep	PPd57-2ms.pep	PPd57-3.pep	PPd57-4ms.pep	PPms-1.pep	PPms-2.pep	PPms-E.pep	PPms-G.pep	Pav5ms.pep	Pavms-2.pep	Pavms-3.pep	Pavms-4.pep	RTms-1.pep	RIms-2.pep	RIMS-5.pep	RIms-6.pep
(SEQ ID NO:88]	H	유	品	B		H	A	H	A I	A	A	e i	A	A	A i	H	H	a	£	A	H	£	A	H	a	8		8	3 5	3 1	9		A	H	e i	A	A 1	8 1	[SEQ ID NO:168]

p	
<b>6</b>	
n	
contin	
1	
H	
0	
O	
3	
a)	
ıre	
digure	
0.0	
ŗŽ,	
H	
•	

TNHNKDYTSVEQCEISITRKPVVA	TNHNKOYTSVEOCEISTTRKPVVA	INHNKDYTSVEOCETSTABKRAN	INHNIOYTSVEOCEISIARNDVVA			TNHNKDYTSVEOCEISIARKPVVA	TNHNKDYTSVEOCEISIARKPVVA	INHNKDYTSVEOCEISIARKPVVA		н	TNHNKDYISVEOCEISIARKPVVA	TNHNKDYISVEOCEISIARKPVVA	TNHNKDYTSVEOCEISIARKPVVA	TWHWKDYTSVEOCEISIARKPVVA	TWHNKDYTSVEOCEISLARKPVVA	TIMENKOYTSVEOCEISIARKPVVA	TIMINKOYTSVEQCEISIARKPVVA	TWHNKDYTSVEOCEISIARKPVVA	TNHNKDYTSVEOCEISIARKPVVA	TNHNKDYTSVEOCEISIARKPVVA	TWHNYDYTSVEOCEISIARKPVVA	TMHNKDYTSVEOCEISIARKPVVA	TWHWKDYTSVEQCEISIARKPVVA	TNHNKDYTSVEQCEISIARKPVVA	TNHNKDYTSVEQCEISIARKPVVA	TNHNKDYTSVEQCEISIARKPVVA	TNHNKDYTSVEQCEISLARKPVVA	TNHNKDYTSVEQCEISIARKFVVA	TINHINKDYISVEQCEISIARKPVVA	TNHNKDYTSVERCEISLARKPVVA	TNHNKDYISVEQCEISIARKPVVA	TNHNKDYISVEQCEISIARKPVVA	TWHNKDYTSVEOCEISIARKPVVA	TNHNKDYTSVEOCEISIARKPVVA	TNHNKDYTSVEOCEISIARKPVVA	TNHNKDYTSVEOCEISIARKPVVA	TNHNKOYTSVEOCEISIARRPVVA	TNENKDYTSVGOCETSTABENEDA	TWHNKOYTSVEOCEISIARKPWA	TNHNKOYTSVEOCEISIARKPVVA	****
Aams-2.pep	Aams-4.pep	Aams-5.pep	Aams-6.pep	Aams-A.pep	Aams-B.pep	Acams-2.pep	Acams-3.pep	Acams-4.pep	Acams-5.pep	Cems-F.pep	Cems-G.pep	Cems-H.pep	Cems-I.pep	LGAms-5.pep	LGAms-6.pep	Mi68Dms.pep	Mims-A.pep	Mims-B.pep	Mims-C.pep	PMms-A.pep	PMms-B.pep	PMms - C. pep	PMms-D.pep	PMms-E.pep	PPd57-1ms.pep	PPd57-2ms.pep	PPd57-3.pep	PPd57-4ms.pep	PPms-1.pep	PPms-2.pep	PPms-E.pep	PPms-G.pep	Pav5ms.pep	Pavms-2.pep	Pavms-3.pep	Pavms-4.pep	RIma-1.pep	RTms-2.pep	RTms-5.pep	RIms-6.pep	****
H		[SEQ ID NO:92]	ü	ij	H	A	A	H	[SEQ ID NO:106]	ü	H	H	A	A	A	A	A	H	H	H	ដ	្ឋ	្អ	8	A.	8	3 1	9 5	3 ;	3 1	A	A	H	a	H	A	H		H	[SEQ ID NO:168]	



### Figure 4

GAAACGITIGCGAAICGITGTICTGCGCTACTIATICTCAAIAIGAGIGTGATCGCTACA GAAACGITIGCGAAICGITGITCTGCGCTACTIAITCTCAAIAIGAGIGIGAICGCIACA	AGIGGGATCGCTACA	AGTGTGATCGCTACA	ATGAGGGTACA	ATGAGTGTGATCGCTACA	ATGAGTGTGATCGCTACA	ATGAGGGTACA	ATGAGGGTACA	AGIGIGATCGCTACA		ATGAGTGTGATCGCTACA	AGTGTGATCGCTACA	AGIGIGALCGCTACA	ATGAGTGTGATCGCTACA	ATGAGTGTGATCGCTACA	AGIGIGALCGCTACA	ATGAGTGTGATCGCTACA	ATGAGTGTGATCGCTACA		ATGAGGGGTACA	ATGAGTGTGATCGCTACA	AGTGTGATCGCTACA	AGTGTGATCGCTACA	AGTGTGATCGCTACA	AGTGTGATCGCTACA		ATGAGTGTGATCGCTACA	AGTGTGATCGCTACA	ATGACTGTGATCGCTACA				AGIGIGATOGCTACA	ATGAGTGTGATCGCTACA	ATGAGGEGATCGCTACA	AGTGTGATCGCTACA	ATGAGTGTGATCGCTACA	ATGAGTGTGAGCGCTACA	AD ALL COLOR OF THE COLOR OF TH	1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-
Aams-2 Aams-4	Aams-5	Aams-6	Aams-A	Aams-B	Acams-2	Acams - 3	Acams-4	Acams-5	Cems-F	Cems-G	Cems-H	Cems-I	LGAMB-5	LGAms-6	Mi68Dms	Mims-A	Mims-B	Mims-C	PMms-A	PMms-B	PMms-C	PMms-D	PMms - E	PPdS7-1ms	PPd57-2ms	PPd57-3	PPd57-4ms	PPms-1	PPms-2	PPMs-E	PPms-G	Pavsms	Pavms-2	Pavms-3	Pavms-4	RTms-1	RIms-2	RTms-5	RTms-6
[SEQ ID NO:87]	B	[SEQ ID NO:93]	[SEQ ID NO:95]	[SEQ ID NO:97]	[SEQ ID NO:99]	[SEQ ID NO:101]	(SEQ ID NO:103)	(SEQ ID NO:105)	(SEQ ID NO:107)	(SEQ ID NO:109]		[SEQ ID NO:113]	[SEQ ID NO:115]	[SEQ ID NO:117]	(SEQ ID NO:119)	H	[SEQ ID NO:123]	(SEQ ID NO:125)	H	(SEQ ID NO:129)	[SEQ ID NO:131]	(SEQ ID NO:133)	[SEQ ID NO:135]	[SEQ ID NO:137]		[SEQ ID NO:141]	[SEQ ID NO:143]		H	[SEQ ID NO:149]	[SEQ ID NO:151]	A	[SEQ ID NO:155]	[SEQ ID NO:157]	[SEQ ID NO:159]	[SEQ ID NO:161]	(SEQ ID NO:163)	a	[SEQ ID NO:167]

96	י ט	U	IJ	ป	บ	ป	ี่	ช	ี่	ช	ี่	ฮ	ฮ	ฮ	ฮ	ฮ	ฮ	ฮ	ฮ	ฮ	ฮ	ð	ð	ឡ	ð	ජි	១	J	ប	ჟ	<u>ქ</u>	ð	បី	ថ	บ	ర	ຽ	ජි
Aams-2	1	Aams-6	Aams-A	Aams-B	Acams-2	Acams-3	Acams-4	Acams-5	Cems-F	Cems-G	Cems-H	Cems-I	LGAms-5	LGAms-6	Mi68Dms	Mims-A	Mims-B	Mims-C	PMms-A	PMms - B	PMms-C	PMms-D	PPd57-1ms	PPd57-2mB	PPd57-3	PPd57-4ms	PPms-1	PPms-2	PPm8-E	PPms-G	Pav5ms	Pavms-2	Pavms-3	Pavms-4	RTms-1	RTms-2	RTms-5	RTms-6
[SEQ ID NO:87]	ID NO:	ED NO:	8.0N	ID NO:9	ID NO:9	ID No:	ID NO:10	ID NO:1	A	ដ	io Oi	ID NO:	[SEQ ID NO:115]	ID NO:	(SEQ ID NO:119)	ID NO:	ID NO:	유	ID NO:	ID NO:	ID NO:	ID NO:	ID NO:13	ID NO:13	ID NO:13	ID NO:1	ID NO:14	A	ID NO:	H	A	A	(SEQ ID NO:155)	H	EQ ID	EQ ID NO:1	_	EQ ID NO:1

aaatgacctacaaggttatatgtcaggcacggtcaatggacactactttgaggtcgaa aaatgacctacaaggtttatatgtcaggcacggtcaatggacactactttgaggtcgaa aaatgacctacaaggittatatgtcaggcacggtcaatggacactactttgaggtcgaa aaatgacctacaaggtttatatgtcaggcacggtcaatggacactactttgaggtcgaa aaatgacctacaaggtttatatgtcaggcacggccaatggacactactttgaggttgaa aaatgacctacaaggittatatgtcaggcacggtcaatggacactactttgaggtcgaa aaatgacctacaaggtttatatgtcgggcacggtcaatggacactactttgaggtcgaa aaatgacctacaaggtttatatgtcaggcacggtcaatggacactactttgaggtcgaa aaaigacctacaaggittatatgtcaggcacggccaatggacactttgaggttgaa aaatgacctacaaggittatatgtcaggcacggccaatggacacttttgaggttgaa aaatgacctacaaggittatatgtcaggcacggtcaatggacactattgaaggtcgaa aaatgacctacaaggttatatgtcaggcacggtcaatggacactactttgaggtcgaa aaatgacctacaaggtttatatgtcaggcacggtcaatggacctactttgaggtcgaa aaatgacctacaaggittatatgtcaggcacggtcaatggacactactttgaggttgaa aaatgacctacaaggittatatgtcaggcacagtcaatggacactagggctcgaa aaatgacctacaaggittatatgtcaggcacggtcgatggacacttactttgaggtcgaa aaatgacctacaaggtttatatgtcaggcacggtcgatggacactactttgaggtcgaa aaatgacctacaaggittatatgtcaggcacggtcgatggacactactttgaggtcgaa aaatgacctacaaggtttatatgtcaggcacggtcgatggacactactttgaggtcgaa aaatgacctacaaggtttatatgccaggcacggtcaatggacactactttcaggttgaa aaatgacctacaaggtttatatgtcaggcacggtcaatggacactacttgaggttgaa aaatgacctacaaggtttatatgtcaggcacggtcaatggacactactttgaggtccaa aaatgacctacaaggtttatatgtcaggcacggtcaatggacactactttgaggttgaa aaatgacctacaaggtttatatgtcaggcacggtcaatggacactactttgaggtcgaa naatgacctacaaggtttatatgtcaggcacggtcaatggacactactttgaggttgaa MATGACCTACAAGGTTTATATGTCAGGCACGGTCAATGGACACTACTTTGAGGTCGAA aatgacctacaaggtttatatgtcaggcacggtcaatggacactactttgaggtcgaa MATGACCTACAAGGTTTATATGTCAGGCACGGTCAATGGACACTACTTTGAGGTTGAA aatgacctacaaggtttatatgtcaggcacggtcaatggacactactttgaggtcgaa <u> patgacctacaaggtttatatgtcaggcacggtcaatggacactactttgaggtcgaa</u> MATGACCTACNAGGTTTATATGTCAGGCACGGTCAATGGACACTACTTTGAGGTCGAA datgacctacaaggtttatatgtcaggcacggtcaatggacactactttgaggtcgaa aatgacctacaaggtttatatgtcaggcacggtcaatggacactactttgaggttgaa aatgacctacaaggttatatgtcaggcacggtcaatggacactactttgaggttgaa AGTGACCTACAAGGTTTATATGTCAGGCACGGTCAATGGACACTACTTTGAGGTTGAA aatgacctacaaggtitatatgtcaggcacggtcaatggacactactttgaggttgaa aatgacctacaaggtttatatgtcaggcacggtcaatggacactactttgaggttgaa aatgacctacaaggttatatatgtcaggcacggtcaatggacactactttgaggttgaa aatgacctacaaggttatatgtcaggcacggtcaatggacactacttgaggtcgaa 

GGCGATGGAAAAGGAAAGCCTTACGAGGGGGAGCAGACGGTAAGGCTGGCT	GGCGATGGAAAAGGAGCCTTACGAGGGGAGCAGACGGTAAGGCTCACTGTCACAAAGGGCACAAAAGGAAAAGCCTTACGAGGGGAGCAGACGGTAAGGCTCACTGTCACAAAGGGAAGCCTTACGAGGGGAGCAAACGGTAAGGCTCACTGTCACAAAGGGAAGCCTTACGAGGGGGAGCAAACGGTAAGGCTCACTGTCACAAAGGGAAGCCTTACGAGGGGAGCAGACGGTAAGGCTCACTGTCACAAAGGCGATGGAAAGGCAAAGCCTTACGAAGGGAGCAGACGGTAAGGCTCACTGTCACAAAGGCGATGAGAAAGCCTTACGAAGGGAGCAGACGGTAAGGCTCACTGTCACAAAGGCCATGAGAAAGCCTTACGAAGGGAGCAGACGGTAAGGCTCACTGTCACAAAGGCCATGAGAAAGCCTTACGAAGGGAGCAGACGGTAAAGCTCACTGTCACAAAGGCCATGAGAAAGCCTTACGAAGGCAGAGGAGCAGATAAGCTCACTGTCACAAAGGCCATGAGAAAGCCTTACGAAAGGCAGAGGAGCAGACGGTAAAGCTCACTGTCACAAAGGCCATAAGGCAAAAGGAAAAGCAAAAGGAAAAGCAAAAGGAAAAGCAAAAGGAAAAGCAAAAGGAAAAGCAAAAGGAAAAGCAAAAGGAAAAGCAAAAGGAAAAGCAAAAGGAAAAGCAAAAGGAAAAGCAAAAGAAAAGCAAAAGAAAAACAAAAGAAAAAA
Aams-2 Aams-4 Aams-5 Aams-6 Aams-Aams-B Acams-3 Acams-2 Acams-3 Acams-3 Acams-4 Acams-2 Cems-4 Cems-4 Cems-6 Cems-H Cems-B Cems-B Cems-A Mins-B Mins-B Mins-B Mins-B PMas-A PMms-A PMms-B PMas-B	Pav5ms Pavms-2 Pavms-4 Pavms-1 RTms-1 RTms-2 RTms-5 Tms-6
	[SEQ ID NO:153] [SEQ ID NO:155] [SEQ ID NO:157] [SEQ ID NO:169] [SEQ ID NO:163] [SEQ ID NO:163] [SEQ ID NO:167]

H	Aams-2	GGCGGACCTCTGCCATTTGCTTGGGATATTTTATCACCACAGTGTCAGTACGGAAGCATA
A	Aams-4	GGCGGACCTCTGCCATTTGCTTGGGATATTTTATCACCACAGTGTCAGTACGGAAGCATA
H	Aams-5	GGCGGACCTCTGCCATTTGCTTGGGATATTTTATCACCACAGTGTCAGTACGGAAGCATA
[SEG ID NO:93]	Aams-6	GGCGGACCICTGCCATTTGCCTGGGATATTTTATCACCACAGTGTTCAGTAACGAAAGAAA
	Aams-A	GGCGGACCTCTGCCATTTGCTTGGGATATTTTATCACCACACACA
ü	Aams-B	GGCGGACCTCTGCCATTTGCGATATTTTATCACTACTACTCACTACTACTACTACTACTAC
H	Acams-2	GGCGGACCTCTGCCATTTGCTTGAGATATTTTATCACCACAGTATCAGTACGGAAGCATA
H	Acams-3	GGCGGACCTCTGCCATTTGCTTGAGATATTTTATCACCACAGTATCAGTACGGAAGCATA
H	Acams-4	GGCGGACCTCTGCCATTTGCTTGGGATATTTTATCACCACAGTATCACTACGGAAGCATA
H	Acams-5	GGCGGACCTCTGCCATTTTGCTTGGGATATTTTATCACCACAGTGTCAGTACGGAAACATA
유	Cems - F	GGCGGACTCTGCCATTTGCTTGGGATATTTTATCACCACAGTGTCACTACGGAAATATA
H	Cems - G	GGCGGACCTCTGCCATTTGCGTATATTTTATCACCACAGTGTCAGTACGGAAACATA
H	Cems-H	GGCGGACCTCTGCCATTTGCCTTGGGATATTTTATCACCACAGTGTCAGTACGGAAACATA
H	Cems-I	GGCGGACCTCTGCCATTTGCTTGGGATATTTTATCACCACAGTGTCAGTACGGAAACATA
ü	LGAms - 5	GGCGGACCTCTGCCATTTGCTTGGGATATTTTATCACCACAGTACGGAAGCATA
H	LGAms-6	GGCGGACCTCTGCCATTTGGGATATTTTGTCACCACAGTATCAGTACGGAAGCATA
H	Мібвршв	GGCGGACCTCTGCCATTTGCTTGGGATATTTTATCACCACAGTGTCAGTACGGAAGCATA
H	Mims-A	GGCGGACCICTGCCATTIGCTTGGGATATTTTATCACCACAGTGTCAGTACGGAAGCATA
ü	Mims-B	GGCGGACCTCTGCCATTTGCTTGGGATATTTTATCACCACAGTGTCAGTACGGAAGCATA
H	Mims-C	GGCGGACCTCTGCCATTTGCTTGGGATATTTTATCACCACAGTGTCAGTACGGAAGCATA
H	PMms-A	GGCGGACCTCTGCCATTTGCTTTGGGATATTCTATCACCACAGAGTCAGTACGGAAGCATA
H	PMms-B	GGCGGACCTCTGCCATTTCCTTGGGATATTCTATCACCACAGAGTCAGTACGGAAGCATA
A	PMms-C	GGCGGACCTCTGCCTTGGGATATTTTATCACCTCAGACTCAGTACGGAAGCATA
H	PMms-D	GGCGGACCTCTGCCATTTGCTTGGGATATTCTATCACCACAGAGTCAGTACGGAAGCATA
H	Mms-E	GGCGGACCTCTGCCATTTGCTTGGGATATTCTATCACCACAGAGTCAGTACGGAAGCATA
a	d57-1ms	GGCGGACCTCTGCCGTTTGCTTGGGATATTTTATCACCACAGACTCAGTACGAAAGCATA
A	Pd57-2mg	GGCGGACCTCTGCCATTTGCTTGGGATATTTTATCACCACAGTGTCAGTACGGAAGCATA
H	PPd57-3	GGCGGACCTCTGCCATTTGCTTGGGATATTTTATCACCACAGTGTCAGTACGGAAAGCATA
Ω	PPd57-4ms	GGCGGACCTCTGCCATTTGCTTGGGATATTCTATCACCACAGAGTCAGTACGGAAGCATA
B	PPms-1	GGCGGACCTCTGCCGTTTGCTTGGGATATTTTATCACCACAGTGTCAGTACGGAAACATA
A	PPms-2	GGCGGACCTCTGCCATTTGCTTGGGATATTTTGTCACCACAGTATCAGTACGGAAGCATA
H	PPm8-E	GGCGGACCTCTGCCATTTGCTTGGGATATTTTATCACCACAGTGCTCAGTACGGAAACATA
A	PPms-G	GGCGGACCTCTGCCATTTGCTTGGGATATTTTATCACCACAGTGTCAGTACGGAAACATA
H	Pav5ms	GGCGGACCTCTGCCATTTGCTTGGGATATTTTATCACCACAGTACAGTACGGAAGCATA
A	Pavms-2	GGCGGACCTCTGCCATTTGCTTGGGATATTTTATCACCACAGTATCAGTACGGAAGCATA
A	Pavms-3	GGCGGACCTCTGCCATTTGGCATATTTTATCACCACAGTATCAGTACGGAAGCATA
R	Pavms-4	GGCGGACCTCTGCCATTTGCTTGGGATATTTTATCACCACAGTATCAGTACGGAAGCATA
H	RTms-1	GGCGGACCTCTGCCATTTGCTTGGGATATTTATCACCACAGTATCAGTACGGAAGCATA
A	RTMs-2	GGCGGACCTCTGCCATTTGCTTGGGATATTTTATCACCACAGTATCAGTACGGAAGCATA
H	RTms-5	GGCGGACCTCTGCCATTTGCTTGGGATATTTTATCACCACAGTGTCAGTACGGAAGCATA
(SEQ ID NO:167)	RTms-6	GGCGGACCCCTGCCATTTGCTTGGGATATTTTATCACCACAGTATCAGTACGGAAGCATA

CCATTCACCAAGTACCCTGAAGACATCCCTGACTATGTAAAGCAGTCATTCCCGGAGGGA CCATTCACCAAGTACCCTGAAGACATCCCTGACTATGTAAAGCAGTCATTCCCGGAGGGA CCATTCACCAAGTACCCTGAAGACATCCCTGACTATGTAAAGCAGTCATTCCCGGAGGGA CCATTCACCAAGTACCCTGAAGACATCCCTGACTATGTAAAGCAGTCATTCCCGGAGGGA CCATTCACCAAGTACCCTGAAGACATCCCTGACTATGTAAAGCAGTCATTCCCGGAGGGA CCATTCACCAAGTACCCTGAAGACATCCCTGACTATGTAAAGCAGTCATTCCCGGAGGGA	CCATTCACCAAGTACCCTGAAGACATCCCTGACTATGTAAAGCAGTCATTCCCGGAAGGA CCATTCACCAAGTACCCTGAAGACATCCCTGACTATGTAAAGCAGTCATTCCCGGAAGGA CCATTCACCAAGTACCCTGAAGACGTCCCTGACTATGTAAAGCAGTCATTCCCGGAGGGA CCATTCACCAAGTACCTGAAGACGTCCCTGACTATGTAAAGCAGTCATTCCCCGGAGGGA CCATTCACCAAGTACCTGAAGACGTCCCTGACTATGTAAAGCAGTCATTCCCGGAGGGA CCATTCACCAAGTACCTGAAGACGTCCCTGACTATGTAAAGCAGTCATTCCCGGAGGGA CCATTCACCAAGTACCCTGAAGACGTCCCTGACTATGTAAAGCAGTCATTCCCGGAGGGA	CCATTCACCAAGTACCCTGAAGACATCCCTGACTATGTAAAGCAGTCATTCCCGGAAGGA CCATTCACCAAGTACCCTGAAGACATCCCTGACTATGTAAAGTAGTCATTCCCGGAAGGA CCATTCACCAAGTACCCTGAAGACATCCCTGACTATGTAAAGTAGTCATTCCCGGAGGGA CCATTCACCAAGTACCCTGAAGACATCCCTGACTATGTAAAGCAGTCATTCCCGGAGGGA CCATTCACCAAGTACCCTGAAGACATCCCTGACTATGTAAAGCAGTCATTCCCGGAGGGA CCATTCACCAAGTACCCTGAAGACATCCCTGACTATGTAAAGCAGTCATTCCCTGAGGGA CCATTCACCAAGTACCCTGAAGACATCCCTGACTATGTAAAGCAGTCATTCCCTGAGGGA CCATTCACCAAGTACCCTGAAGACATCCCTGACTATGTAAAACAGTCATTCCCTGAGGGA CCATTCACCAAGTACCCTGAAGACATCCCTGACTATGTAAAACAGTCATTCCCTGAGGGA CCATTCACCAAGTACCCTGAAGACATTCCCTGAGGGA	CCATTCACCAAGTACCTGAAGACATCCCTGACTATGTAAAGCAGTCATTCCCTGAGGGA CCATTCACCAAGTACCTGAAGACATTCCTGACTATGTAAAACAGTCATTCCCTGAGGGA CCATTCACCAAGTACCTGAAGACATTCCTGACTATGTAAAACAGTCATTCCCTGAGGGA CCATTCACCAAGTACCTGAAGACATCCCTGACTATGTAAAGCAGTCATTCCCCGGAGGGA CCATTCACCAAGTACCTTGAAGACATCCCTGACTATGTAAAGCAGTCATTCCCCGGAGGGA CCATTCACCAAGTACCCTGAAGACATCCCTGACTATGTAAAGCAGTCATTCCCCGGAGGGA CCATTCACCAAGTACCCTGAAGACATCCCTGACTATGTAAAGCAGTCATTCCCGGAGGGA CCATTCACCAAGTACCCTGAAGACATCCCTGACTATGTAAAGCAGTCATTCCCGGAAGGA CCATTCACCAAGTACCCTGAAGACATCCCTGACTATGTAAAGCAGTCATTCCCGGAAGGA CCATTCACCAAGTACCCTGAAGACATCCCTGACTATGTAAAGCAGTCATTCCCGGAAGGA CCATTCACCAAGTACCCTGAAGACATCCCTGACTATGTAAAGCAGTCATTCCCGGAAGGA CCATTCACCAAGTACCCTGAAGACATCCCTGACTATGTAAAAGCAGTCATTCCCGGAAGGA CCATTCACCAAGTACCCTGAAGACATCCCTGACTATGTAAAAGCAGTCATTCCCGGAAGGA CCATTCACCAAGTACCCTGAAGACATCCCTGACTATGTAAAAGCAGTCATTCCCGGAAGGA CCATTCACCAAGTACCCTGAAGACATCCCTGACTATGTAAAAGCAGTCATTCCCGGAAGGA CCATTCACCAAGTACCCTGAAGACATCCCTGACTATGTAAAAGCAGTCATTCCCGGAAGGA CCATTCACCAAGTACCCTGAAGACATCCCTGACTATGTAAAAGCAGTCATTCCCGGAAGGA CCATTCACCAAGTACCCTGAAGACATTCTTAAAAACACTCATTCCCGGAAGGA CCATTCACCAAGTACCCTGAACACTATTGTAAAAGCAGTCATTCCCGGAAGGA CCATTCACCAAGTACCCTGAACACTATTGTAAAAGCAGTCATTCCCGGAAGGA CCATTCACCAAGTACCCTGAACACCTGACTATTGTAAAAGCAGTCATTCCCGGAAGGA CCATTCACCAAGTACCCTGAACACTATTGTAAAAGCAGTCATTCCCGGAAGGA CCATTCACCAAGTACCCTGAACACACTATTGTAAAAGCAGTCATTCCCGGAAGGA CCATTCACCAAGAACACCCTGACTATTGTAAAAACATATCCCGGAAGGA CCATTCACCAAGAACACCCTGACTATTGTAAAACACTCATTCCCGGAAGGA CCATTCACCAAGTACCCTGAACAACTATTGTAAAACATATTCCCGGAAGGA CCATTCACAAGTACCCTGAACAACTATTGTAAAACATATCCCTGAACGAAC
Aams-2 Aams-4 Aams-5 Aams-6 Aams-A Aams-B Acams-2	Acams - 4 Acams - 5 Cems - 7 Cems - 7 Cems - 1 Cems - 1	LGAMES-5 LGAMES-6 MiGS-6 Mins-A Mims-A Mims-C PMms-A PMms-B PMms-C PMms-C PMms-C	PMms-E PPd57-1ms PPd57-2ms PPd57-3 PPd57-4ms PPms-1 PPms-2 PPms-G PPms-G Pavfms Pavms-3 Pavms-3 Pavms-1 RTms-1 RTms-5 RTms-5
6666666	[SEQ ID NO:101] [SEQ ID NO:103] [SEQ ID NO:105] [SEQ ID NO:107] [SEQ ID NO:111] [SEQ ID NO:1113]	200000000000000000000000000000000000000	[SEQ ID NO:135] [SEQ ID NO:137] [SEQ ID NO:141] [SEQ ID NO:141] [SEQ ID NO:143] [SEQ ID NO:144] [SEQ ID NO:144] [SEQ ID NO:151] [SEQ ID NO:151] [SEQ ID NO:153] [SEQ ID NO:153] [SEQ ID NO:154] [SEQ ID NO:157]

SEQUENCY OF SEQUEN

ICCAGCATCCAAGGCAACTGTTTCATCTACCATGTCAAGTTCTCTGGTTTGAACTTTCCT ICCAGCATCCAAGGCAACTGTTCATCTACCATGTCAAGTTCTGGTTTGAACTTTCCT TCCAGCATCCAAGGCAACTGTTTCATCTACCATGTCAAGTTCTCTGGTTTGAACTTTCCT TCCAGCATCCAAGGCAACTGTTCATCTACCATGTCAAGTTCTGGGTTTGAACTTTCCT TCCAGCATCCAAGGCAACTGTTTCATCTACCATGTCAAGTTCTGGGTTTGAACTTTCCT ICCAGCATCCAAGGCAACTGTTTCATCTACCATGTCAAGTTCTCTGGTTTGAACTTTCCT ICCAGCAICCAAGGCAACTGTTTCATCCACCAIGTCAAGTTCTGGGTTTGAACTTTCCT TCCAGCATCCAAGGCAACTGTTCACCTACCATGTCAAGTTCTCTGGTTTGAACTTTCCT ICCAGCATCCAAGGCAACTGTTTCACCTACCATGTCAAGTTCTCTGGTTTGAACTTTCCT TCCAGCATCCAAGGCAACTGTTTCACCTACCATGTCAAGTTCTCTGGTTTGAACTTTCCT ICCAGCATCCAAGGCAACTGTTTCACCTACCATGTCAAGTTCTCTGGTTTGAACTTTCCT TCCAGCATCCAAGGCAACTGTTTCACCTACCATGTCAAGTTCTCTGGTTTGAACTTTCCT ICCAGCATCCAAGGCAACTGTTCATCTACCATGTCAAGTTCTCTGGTTTGAACTTTCCT ICCAGCATCCAAGGCAACTGTTCATCTACCATGTCAAGTTCTGTGGTTTGAACTTTCCT TCCAGCATCCAAGGCAACTGTTTCACCTACCATGTCAAGTTCTCTGGTTTGAACTTTCCT rcaacaiccaaggaacigiiiicacciaccaigicaagiiciciggiiigaaciiiicc rccagcatccaaggcaactgttracctagccatgtcaagttctggtttgaactttgct ICCAGCATCCAAGGCAACTGTTTCACCTACCATGTCAAGTTCTCTGGTTTGAACTTTCCT TCCAGCATCCAAGGTAACTGTTTCATCTACAATGTCAAGTTCTGGGTTTGAACTTTCCT ICCAGCATCCAAGGTAACTGTTCATCTACAATGTCAAGTTCTCTGGTTTGAACTTTCCT ICCAGCATCCAAGGCAACTGTTTCATCTACAATGTCAAGTTCTCTGGTTTGAACTTTCCT TCCAGCATCCAAGGCAACTGTTTCATCTACAATGTCAAGTTCTCTGGTTTTGAACTTTCCT ICCAGCATCCAAGGCAACTGTTTCATCTACCATGTCAAGTTCTCTGGTTTGAACTTTCCI ICCAGCATGCAAGGCAACTGTTTCATCTACAATGTCAAGTTCTCTGGTTTGAACTTTCCT TCCAGCATGCAAGGCAACTGTTTCATCTACAATGTCAAGTTCTCTGGTTTTGAACTTTCCT ICCAGCATCCAAGGCAACTGTTTCATCTACCATGTCAAGTTCTCTGGTTTGAACTTTCCT TCCAGCATCCAAGGCAACTGTTTCATCTACAATGTCAAGTTCTCTGGTTTTGAACTTTCCT ICCAGCAICCAAGGCAACIGITICACCIACCAIGICAAGITCICIGGIITIGAACITITCCI **ICCAGCAICCAAGGCAACTGITICAICIACCAIGICAAGIICCIGGIIIIGAACIIIICCI** TCCAGCATCCAAGGCAACTGTTTCATCTACCATGTCAAGTTTTCTGGTTTTGAACTTTTCCT TOCAGCATCCAAGGCAACTGTTCATCTACCATGTCAAGTTCTCTGGTTTTGAACTTTCCT TCCAGCATCCAAGGCAACTGTTCATCTACCATGTCAAGTTCTGGGTTTGAACTTTCCT TCCAGCATCCAAGGCAACTGTTTCACCTACCATGTCAAGTTCTCTGGTTTGAACTTTCCT ICCAGCATCCAAGGCAACTGTTTCATCTACCATGTCAAGTTCTCTGGTTTGAACTTTCCT TCCAGCATCCAAGGCAACTGTTCACCTACCATGTCAAGTTCTGGGTTTGAACTTTCCT TCCAGCATCCAAGGCAACTGTTCATCTACCATGTCAAGTTTTTCTGGGTTTGAACTTTTCCT TCCAGCATCCAAGGCAACTGTTTCATCTACCATGTCAAGTTCTCTGGTTTGAACTTTCCT ICCAGCAICCAAGGCAACIGITIICACCIACCAIGICAAGIICICIGGIIIIGAACIIIICC ICCAGCAICCAAGGCAACTGITTCAICTACCAIGICAAGITCICIGGITTGAACTTTCCI ICCAGCATCCAAGGCAACTGTTCATCTACCATGTCAAGTTCTGGGTTTGAACTTTCCT TCCAGCATCCAAGGCAACTGTTTCATCTACCATGTCAAGTTCTCTGGGTTTGAACTTT PPd57-2ms PPd57-1mg PPd57-4ms LGAms-6 1cams-2 GAMS-5 Mi68Dms PPd57-3 Acams-3 Acams-4 Acams-5 Cems-H Mims-A Mims-B Cems-G Mims-C PMms-A PMms-B Pavms-2 Pavms-3 Parms-4 lams-A Jame-B Aams-5 lams-6 Cems-F PMms - C PMms-D PMms-E PPms-2 PPms-E PPms-G PPms-1 Pav5ms Tms-2 Tms-5 3-smr RTms-1 NO:107 NO:111 NO:101 NO:103 NO: 105 NO:109 NO:113 NO:115 NO:117 NO:119 NO:121 NO:135] NO:139] NO:145] NO:123 NO:125 NO:127 NO: 129 NO: 131 NO:133] NO: 137 NO:141 NO:149 NO:151 NO:95] NO:143 NO:147 No: 153 NO:159 NO:161 NO:163 NO: 99 888888888 88888 H A A B A H H A 888 Η ü H H H ü SEQ SEQ SEO SEQ SEO

CCCAATGGACCTGTTATGCAGAAGAAGACACAGGGCTGGGGAACCCCACTCTGAGCGTCTC CCCAATGGACCTGTTATGCAGAAGAAGACACAGGGCTGGGGAACCCCACTCTGAGCGTCTC CCCAATGGACCTGTTATGCAGAAGAAGACACAGGGCTGGGAAACCCCACTCTGAGCGTCTC CCCAATGGACCTGTTATGCAGAAGAAGACACAGGGCTGGGAACCCCACTCTGAGCGTCTC CCCAATGGACCTGTTATGCAGAAGAAGACACAGGGCTGGGAACCCCACTCTGAGCGTCTC CCCAATGGACCTGTTATGCAGAAGAAGACACAGGGCTGGGAACCCCACTCTGAGCGTCTC CCCAATGGACCTGTAATGCAGAAGAACACACAGGGCTGGGAACCCCACTCTGAGCGTCTC CCCAATGGACCTGTGATGCAGAAGAACACAGGGCTGGGAACCCCACTCTGAGCGTCTC CCCAATGGACCTGTGATGCAGAAGAACACAGGGCTGGGGAACCCCACTCTGAGCGTCTC CCCAATGGACCTGTGATGCAGAAGAACACAGGGCTGGGGAACCCCACTCTGAGCGTCTC CCCAATGGACCTGTGATGCAGAAGAACACAGGGCTGGGAACCCCACTCTGAGCGTCTC CCCAATGGACCTGTGATGCAGAAGAACACACAGGGCTGGGAACCCCACTCTGAGCGTCTC CCCAATGGACCTGTGATGCAGAAGAACACAGGGCTGGGAACCCCACTCTGAGCGTCTC CCCAATGGACCTGTGATGCAGAAGAACACACGGGCTGGGAACCCCACTCTGAGCGTCTC CCCAATGGACCTGTGATGCAGAAGAACACACAGGGCTGGGAACCCCACTCTGAGCGTCTC CCCAATGGACCTGTGATGCAGAAGAACACACAGGGCTGGGAACCCCACTCTGAGCGTCTC CCCAATGGACCTGTGATGCAAAGAAACACACAGGGCTGGGAACCCCACTCTGAGCGTCTC CCCAATGGACCTGTGATGCAAAGAAAACACACAGGGCTGGGAACCCCACTCTGAGCGTTCC CCCAATGGACCTGTGATGCAAAAAAAACACACAGGGCTGGGAACCCCACTCTGAGCGTTCC CCCAATGGACCTGTGATGCAAAAAAAACACACAGGGCTGGGAACCCCACTCTGAGCGTTCC CCCAATGGACCTGTGATGCAAAAAAAACACACAGGGCTGGGAACCCCACTCTGAGCGTTCC CCCAATGGACCTGTGATGCAAAAAAAACACACAGGGCTGGGAACCCCACTCTGAGCGTTCC CCCAATGGACCTGTGATGCAAAAGACACACAGGGCTGGGAACCCCACTCTGAGCGTTCC CCCAATGGACCTGTGATGCAAAAAAACACACAGGGCTGGGAACCCCACTCTGAGCGTTCC CCCAATGGACCTGTGATGCAAAAAAACACACAGGGCTGGGAACCCCCACTCTGAGCGTTCC CCCAATGGACCTGTGATGCAAAAGACACACACAGGGCTGGGAACCCCACTCTGAGCGTTCC CCCAATGGACCTGTGATGCAAAAAACACACACACACACAC	CCCAATGGACCTGTTATGCAAAAGAAGACACAGGGTTGGGAACCCAACACTGAGCGTCTTCCCCAATGGACCTGAGGTTGGGAACCCCAACACTGAGCGTCTTCCCCAATGGACCCTAAAAGAAGAAGACACAGGGGTTGGGAACCCCAACACTGAGCGTCTTCCCCCAATGGACCTGTTATGCAAAAGAAGACACAGGGCTTGGGAACCCCAACACTGAGCGTCTTCCCCCAATGGACCTGTTATGCAAAAGAAGACACAGGGCTGGGAACCCCAACACTGAGCGTCTCCCCCAATGGACCTGTTATGCAAAGAAGACACAGGGCTGGGAACCCCAACTGAGCGTCTCCCCCAATGGACCTGTTATGCAAAGAAGACACAGGGCTGGGAACCCCAACTGAGCGTCTCCCCCAATGGACCTGTTATGCAAAGAAGACACAGGGCTGGGAACCCCAACTGAGCGTCTCCCCCAATGGACCTGTTATGCAAAGAAGACACAGGGCTGGGAACCCCAACTGAGCGTCTCCCCCAATGGACCTGTTATGCAAAGAAGACACAGGGCTGGGAACCCCAACACTGAGCGTCTCCCCCAATGGACCTGTTATGCAAAGAAGACACAGGGCTGGGAACCCCAACTGAGCGTCTCCCCCAATGGACCTTGTATGCAAGAAGACACAGGGCTGGGAACCCCAACTGAGCGTCTCCCCCAATGGACCTTGTATGCAAAGAAGACACAGGGCTGGGAACCCCAACTGAGCGTCTCCCCCAATGGACCTTGTATGCAGAAGAAGACACAGGGCTGGGAACCCCAACTGAGCGTCTCCCCCAATGGACCTTGTATGCAGAAGAAGACACAGGGCTGGGAACCCCAACTGAGCGTCTCCCCCAATGGACCTTGAGAGACACACAGGGATCTCCCCCAATGGACCTTGAGAGACACACAGGGATCTCCCCCAATGGACCTTGAGAAGAAAAAAAA
Aams-2 Aams-4 Aams-5 Aams-6 Aams-A Aams-A Acams-2 Acams-3 Acams-4 Acams-4 Acams-7 Cems-R Cems-R Cems-R Cems-I LGAms-5 LGAms-6 MiG8Dms Mims-A Mims-A Mims-A	PMms-B PMms-C PMms-C PMms-C PMms-D PPd57-1ms PPd57-3 PPd57-3 PPd57-3 PPms-1 PPms-Z PRTms-1 RTms-1 RTms-5 RTms-5
[SEQ ID NO:87] [SEQ ID NO:88] [SEQ ID NO:91] [SEQ ID NO:93] [SEQ ID NO:95] [SEQ ID NO:99] [SEQ ID NO:101] [SEQ ID NO:101] [SEQ ID NO:104] [SEQ ID NO:105] [SEQ ID NO:115] [SEQ ID NO:115] [SEQ ID NO:115] [SEQ ID NO:115] [SEQ ID NO:121] [SEQ ID NO:121] [SEQ ID NO:121] [SEQ ID NO:123] [SEQ ID NO:123] [SEQ ID NO:123] [SEQ ID NO:123]	

:ataggaaacaactitatggctctgaagttagaaggaggc htaggaaactttatggctctgaagttagaaggagg iataggaacaactittaiggcictgaagttagaaggagg htaggaaacaactittatggctctgaagttagaaggaggc htaggaacaactttatggctctgaagttagaaggagg ataggaaacaactittatggctctgaagttagaaggagg ataggaacaactitatggctctgaagttagaaggagg htaggaacaactitatggctctgaagttagaaggaggc ataggaaacaactttgtggctctgaagttagaaggaggc 'ATAGGAAACAACTITIATGGCTCTGAAGTTAGAAGGAGGC ATAGGAAACAACTTTATGGCTCTGAAGTTAGAAGGAGGC ataggaaacaactttatggctctgaagttagaaggaggt ATAGGAAACAACTTTATGGCTCTGAAGTTAGAAGGAGGC CTAGGAAACAACTTTATGGCTCTGAAGTTAGAAGGAGGT ATAGGAAACAACTTTATGGCTCTGAAGTTAGAAGGAGGC ATAGGAAACAACTTTATGGCTCTGAAGTTAGAAGGAGGC ATAGGAAACAACTTTATGGCTCTGAAGTTAGGAGGAGGC ATAGGAAACAACTTTATGGCTCTGAAGTTAGAAGGAGGC ATAGGAAACAACTTTATGGCTCTGAAGTTGGAAGGAGGT

*I*ATAGGAAACAACTTTATGGCTCTGAAGTTAGAAGGAGG

*WTAGGAAACAACTTTATGGCTCTGAAGTTAGAAGGAGGC* 

H	Aams-2	TITGCACGAGATGGAATGCTGATAGGAAACAACTTTATGGCTCTGAAGTI
H	Aams-4	TITGCACGAGATGCAGATACGAAACAACTTTATGGCTCTGAAGTT
H	Aams-5	TITGCACGAGACGGAATGCTGATAGGAAACAACTTTATGGCTCTGAAGTT
H	Aams-6	TITGCACGAGACGGAATGCTGATAGGAAACAACTTTATGGCTCTGAAGTT
H	Aams-A	TITIGCACGAGATIGGAATGCTGATAGGAAACAACTITTATGGCTCTGAAGTT
H	Aams-B	TITGCACGAGATGGAATGCTGATAGGAAACAACTTTATGGCTCTGAAGTT
H	Acams-2	TITGCACGAGATGCTGATAGGAAACAACTITATGGCTCTGAAGTT
A:	Acams-3	TITGCACGAGATGCAATGCTGATAGGAAACAACTTTATGGCTCTGAAGTT
A i	Acams-4	TITGCACGAGATGGAATGCTGATAGGAAACAACTITATGGCTCTGAAGTT
ដ	Acams-5	TTTGCACGGGGTGGAATGCTGATAGGAAACAACTTTGTGGGCTCTGAAGTT
A I	Cems-F	TITGCACGGGGTGGAATGCTGATAGGAAACAACTITATGGCTCTGAAGTT
H	Cems-G	CITGCACGGGTGGAATGCTGATAGGAAACAACTTTATGGCTCTGAAGTT
H	Cems-H	TITGCACGGGTGGAATGCTGATAGGAAACAACTTTATGGCTCTGAAGTT
A	Cems-I	TITGCACGGGGTGGAATGCTGATAGGAAACAACTTTATGGCTCTGAAGTT
A	LGAms - 5	TTAGCACGAGATGGAATGCTGATAGGAAACAACTTTATGGCTCTGAAGTT
H	LGAms-6	TITGCACGAGATGCAGCTAGGAAACAACITTATGGCTCTGAAGTT
H	Mi68Dms	TITGCACGGGTGGAATGCTGATAGGAAACAACTTTATGGCTCTGAAGTT
A	Mims-A	TITGCACGGGTGGAATGCTGATAGGAAACAACTITATGGCTCTGAAGTT
A	Mims - B	TITGCACGGGTGGAATGCTGATAGGAAACAACTITTATGGCTCTGAAGTT
H	Mims-C	TITGCACGGGGTGGAATGCTGATAGGAAACAACTTTATGGCTCTGAAGTT
H	PMms-A	TATGCACAGATGGAATGCTGATAGGAAACAACTTTATGGCTCTGAAGTT
H	PMms - B	TTTGCACGAGATGGAATGCTGATAGGAAACAACTTTATGGCTCTGAAGTT
H	PMms - C	TATGCACGAGATGGAATGCTGATAGGAAACAACTTTTATGGCTCTGAAGTT
A i	D-SmMd	TITGCACGAGAIGGAAIGCIGAIAGGAAACAACTIITAIGGCICIGAAGIC
A	PMms-E	TITGCACGAGATGGAATGCTGATAGGAAACAACTTTATGGCTCTGAAGTT
H	PPd57-1ms	TATGCACGAGATGGAATGCTGATAGGAAACAACTTTATGGCTCTGAAGTT
H	PPd57-2ms	TITGCACGAGATGGAATGCTGATAGGAAACAACTTTATGGCTCTGAAGTT
A	PPd57-3	TITGCACGAGATGGAATGCTGATAGGAAACAACTTTATGGCTCTGAAGTT
A	PPd57-4ms	TITGCACGAGATGGAATGCTGATAGGAAACAACTTTATGGCTCTGAAGTT
A	PPms-1	TITGCACGGGGGGAATGCTGATAGGAAACAACTTTATGGCTCTGAAGTT
A	PPms-2	CITGCACGAGATGCAGCTGCTAGGAAACAACTTTATGGCTCTGAAGTT
A	PPms - E	TTTGCACGGGGTGGAATGCTGATAGGAAACAACTTTATGGCTCTGAAGTT
A	PPms-G	TITGCACGGGGTGGAATGCTGATAGGAAACAACTTTATGGCTCTGAAGTT
A	Pav5ms	TITGCACGAGAIGGAAIGCIGAIAGGAAACAACITIAIGGCICIGAAGII
A	Pavms-2	TITGCACGAGATGGATTGCTGATAGGAAACAACTTTATGGCTCTGAAGTT
H	Pavms-3	TTTGCACGAGATGGAATGCTGATAGGAAACAACTTTATGGCTCTGAAGTT
A	Pavms-4	TITGCACGAGAIGGAAIGCIGAIAGGAAACAACIIITAIGGCICIGAAGII
B	RTms-1	TITGCACGAGATGGAATGCTGATAGGAAACAACTITATGGCTCTGAAGTT
A I	RTms-2	TITGCACGAGATGGAATGCTGATAGGAAACAACTITATGGCTCTGAAGTT
<b>A</b> !	RTms-5	TTTGCACGGGGTGGAATGCTGATAGGAAACAACTTTATGGCTCTGAAGTT
SEQ ID NO:167	RTms-6	TITGCACGAGATGCATACGAAACAACTITATGGCTCTGAAGTT
		化水水水水 玩 化二苯甲基苯甲基苯甲基苯甲基苯甲基苯甲基苯甲基苯甲基苯甲基苯甲基苯甲基苯甲基苯甲基苯

ataggaaactttatggctctgaagttggaagggg **ATAGGAAACAACTTTATGGCTCTGAAGTTGGAAGGAGGT** ATAGGAAACAACTTTATGGCTCTGAAGTTGGAAGGAGGT ATAGGAAACAACTTTATGGCTCTGAAGTTGGAAGGAGGT ataggaaactittatggctctgaagtcggaaggaggt

ataggaaactttatggctctgaagttagaagggc ataggaaacaactttatggctctgaagttagaaggaggc ataggaaacatttatggctctgaagttggaagga ataggaaacaacittaiggcicigaagitagaaggagg CTAGGAAACAACTTTATGGCTCTGAAGTTAGAAGGAGGT ataggaaacaactttatggctctgaagttagaagagg ataggaacaactttatggctctgaagttagaaggagg **ATAGGAAACAACTTTATGGCTCTGAAGTTAGAAGGAGGC** ataggaaacaactitatggctctgaagttagaagaggc ataggaaactittatggctctgaagttagaaggagg **ATAGGAAACAACTITTATGGCTCTGAAGITTAGAAGGAGGC** ataggaaacaactttatggctctgaagttagaaggaggc **4TAGGAAACAACTTTATGGCTCTGAAGTTAGAAGGAGGC LTAGGAAACAACTTTATGGCTCTGAAGTTAGAAGGAGGC** ataggaaacaactittatggctctgaagttagaaggagg

CAAGGCAAAGAAGCCTGTGAAGATGCCA CAAGGCAAAGAAGCCTGTGAAGATGCCA CAAGGCAAAGAAGCCTGTGAAGATGCCA CAAGGCAAAGAAGCCTGTGAAGATGCCA CAAGGCAAAGCCTGTGAAGATGCCA CAAGGCAAAGACCTGTGAAGATGCCA CAAGGCAAAGAAGCCTGTGAAGATGCCA CAAGGCAAAGAAGCCTGTGAAGATGCCA CAAGGCAAAGAAACCTGTGAAGATGCCA CAAGGCAAAGAAGCCTGTGAAGATGCCA CAAGGCAAAGAAGCCTGTGAAGATGCCA

CAAGGCAAAGAAGCCTGTGAAGATGCCA

CAAGGCAAAGAAGCCTGTGATGATGCCA CAAGGCAAAGAAGCCTGTGATGATGCCA CAAGGCAAAGAAGCCTGTGAAGATGCCA CAAGGCAAAGAAGCCTGTGAAGATGCCA

CAAGGCAAAGAAGCCTGTGATGATGCT

CAAGGCAAAGAAGCCTGTGAAGATGCCA CAAGGCAAAGAAGCCCGTGAAGATGCCA CAAGGCAAAGAAGCCTGTGAAGATGCCA CAAGGCAAAGAAGACTGTGAAGATGCCA CAAGGCAAAGAACACTGCGAAGATGCCA CAAGGCAAAGAAGACTGTGAAGATGCCA CAAGGCAAAGAAGACTGTGAAGATGCCA CAAGGCAAAGAAGCCTGTGAAGATGCCA CAAGGCAAAGAAGCCTGTGAAGATGCCA CAAGGCAAAGAAGCCTGTGAAGATGCCA

CAAGGCAAAGAAGCCTGTGATGATGCCCA

CAAGGCAAAGAAGCCTGTGAAGATGCCA

CAAGGCAAAGAGCCTGTGAAGATGCCA CAGGGCAAAGAGCCTGTGAAGATGCCA

CAAGGCAAAGAAGCCTGTGAAGATGCCA CAAGGCAAAGAAGCCTGTGATGATGCCT CAAGGCAAAGAAGCCTGTGATGATGCCA CAAGGCAAAGAAGCCTGTGATGATGCCT

CAAGGCAAAGAAGCTTGTGAAGATGCCA CAAGGCAAAGAAGCCTGTGAAGATGCCA CAAGGCAAAGAAGCCTGTGAAGATGCCA CAAGGCAAAGAAGCCTGTGAAGATGCCA CAAGGCAAAGAAGCCTGTGAAGATGCCA

### igure 4 continued

经存在存在存在存在 安京

3GGTATCATTATGTTGACCGCAAACTGGATGTAACCAATCACAACAAGGATTACACTTCC gggtatcattatgttgaccgcaactggatgtaaccaatcacaacaaggattacacttcc GGGTATCATTATGTTGACCGCAAACTGGATGTAACCAATCACAACAAGGATTACACTTCC GGGIATCATTATGTTGACCGCAAACTGGATGTAACCAATCACAACAAGGATTACACTTCC 3GGTATCACTATGTTGACCGCAAACTGGATGTAACCAATCACAACAAGGATTACACTTCC GGGTATCACTATGTTGACCGCAAACTGGATGTAACCAATCACAAGAATTACACTTTCC GGGTATCACTATGTTGACCGCAAACTGGATGTAACCAATCACAACAAGGATTACACTTCC gggtatcattatgttgaccgcaaactggatgtaaccaatcaacaaggattacatttcc GGGTATCATTATGTTGACCGCAAACTGGATGTAACCAATCACAACAAGAATTACATTTCC gggtatcattatgttgaccgcaaactggatgtaaccaatcacaacaaggattacatttcc GGGTATCATTATGTTGACCGCAAACTGGATGTAACCAATCACAAGAATTACATTTCC GGGTATCACTATGCTGACCGCAAACTGGATGTAACCAATCACAACAAGGATTACACCTCC GGGTATCACTATGTTGACCGCAAACTGGATGTAACCAATCACAACAAGGATTACACTTCC GGGTATCACTATGTTGACCGCAAACTGGATGTAACCAATCACAACAAGGATTACACTTCC GGGIAICAITAIGITGACCGCAAACTGGAIGIAACCAAICACAACAAGGATIACACTICC GGGIAICAITAIGITGACCGCAAACTGGATGIAACCAAITGACAACAAGGATIACACITCC GGGTATCATTATGTTGACCGCAAACTGGATGTAACCAATCACAACAAGGATTACACTTCC GGGTATCATTATGTTGACCGCAAACTGGATGTAACCAATCACAACAAGGATTACACTTCC ggatatcactatgttgaccgcaaattggatgtaaccaatcacaacaaggattacacttcc GGGIAICACTAIGITGACCGCAAATTGGATGIAACCAAICACAACAAGGATTACACTTCC ggatatcactatgttgaccgcaaattggatgtaaccaatcacaacaaggattacacttcc GGGIATCACTATGTTGACCGCAAATTGGATGTAACCAATCACAACAAGGATTACACTTCC ggatatcactatgttgaccgcaaattggatgtaaccaatcaaacaaggattacacttcc gggtatcattatgttgaccgcaaactggatgtaaccaatcacaacaaggattacacttcc gggtatcactatgttgaccgcaaattggatgtaaccaatcacaacaagaattacacttcc GGGTATCATTATGTTGACCGCAAACTGGATGTAACCAATCACAACAAGGATTACACTTCC GGGIATCACTATGTTGACCGCAAATTGGATGTAACCAATCACAACAAGGATTACACTTCC GGGTATCATTATGTTGACCGCAAACTGGATGTAACCAATCACAACAAGGATTACATTTCC GGGTATCACTATGTTGACCGCAAACTGGATGTAACCAATCACAACAAGGATTACACTTCC GGGIAICATTAIGITGACCGCAAACTGGAIGTAACCAAICAACAACAAGAATTACAITICC GGGTATCATTATGTTGACCGCAAACTGGATGTAACCAATCACAACAAGAATTACATTTCC **GGGTATCACTATGTTGACCGCAAACTGGATGTAACCAATCACAACAAGGATTACACTTCC** GGGTATCACTATGTTGACCGCAAACTGGATGTAACCAATCACAACAAGGATTACACTTCC GGGIATCACTAIGTIGACCGCAAACTGGAIGTAACCAATCACAACAAGGATTACACTICC 3GGTATCACTATGTTGACCGCAAACTGGTTGTAACCAATCACAACAAGGATTACACTTCC 3GGTATCACTATGTTGACCGCAAACTGGATGTAACCAATCACAACAAGGATTACACTTCC GGGTATCACTATGTTGACCGCAAACTGGATGTAACCAATCACAACAAGGATTACACTTCC GGGTATCATTATGTTGACCGCAAACTGGATGTAACCAATCACAACAAGGATTACACTTCC 3GGTATCACTATGTTGACCGCAAACTGGATGTAACCAATCACAACAAGGATTACACCTCC 化化物化物物化物化物化物化物化物化物 医阿拉伯氏 医阿拉 医水杨二氏试验检尿 医水杨素 医医水杨素 PPd57-1ms Pd57-2ms PPdS7-4ms Acams-2 LGAms-6 Acams-3 Acams-4 Acams-5 GAMB-5 M168Dms Mims-A Mims-B Pd57-3 Aams-6 Aams-A Aams-B Cema-G Cema-F Cems-H Pavms-2 Jems-I Mims-C PMms-A PMma - B PMmg-C PPms-1 PPms-2 PPms-E Pavms-3 Pavms-4 O-Smire-D PMme-E PPma-G Pav5ms RTms-2 T-BMIX 2Tm8-5 1mg-6 NO:97] NO:101 NO:103 NO: 105 NO:107 NO:109 NO:113] NO:115 NO:119] NO:135] NO: 111 NO:117 NO:121 NO:93] NO:95] NO:99] NO: 123 NO:125 NO:127 NO:129 NO:131 NO:133 NO:139 NO:145] NO:137 NO:141 NO:143 NO:147 NO:149] NO:151 NO:153] NO:155 NO:157 NO:159 NO:161] NO:165 NO:167 8888888888888 H H 88 H H Н Н H 2222222222 88888 H 038] 038] 

34/94

C - CHE	N I I I I I I	TO THE PERSON OF			A-mark A-mark	Acans - 2		Acams-4	Acams-5	Cems-F	Cema-G	Cems-H	Cema-I	LGAms-5	LGAms-6	Mi68Dms	Mims-A	Mims-B	Mims-C	PMms-A	PMms-B	PMms-C	D-smmd	PMms - E	PPdS7-1ms	PPd57-2ms	PPG57-3	Draw 1			D D D D D D D D D D D D D D D D D D D			Pavms-3	Pavms-4	RTms-1	RTms-2	RIms-5	RIms-6	经存储的现在分词 医红红 按 我 化环状状状状状状状 计存储设计 大學 化水杨素
(SEO ID NO:87)	£	1	96	F	H	a	B	ΩĦ		A	H	ID NO.111]	ID NO:113]	ID NO:115]	ID NO:117]	1D NO:119]	ID NO.121	LD NO:123]	LD NO:125]	ID NO:127]	ID NO:129]	ID NO:131	1D NO:133	ID NO:135]		TO NO. 139	NO: 141]	TD MO:145]	TD NO.147]	TD NO.149]	ID NO: 151	ID NO:153]	ID NO:155]	[SEQ ID NO:157]	ID NO:159]	ID NO:161]	ID NO:163]	ID NO:165]	[SEQ ID NO:167] R	

SVIAKOMTASTLS*QOEARRRLOSRTRERDTWDLSREQISRROMEPGMTRYLWLALLKKP SVIAKOMTASTLS*QOEARRRLOSRTRERDTWDLSREQISRROMEPGMTRYLWLALLKKP SVIAKOMTASTLS*QOEARRLOSRTRERDTWDLSREQISRROMEPGMTRYLWLALLKKP SVIAKOMTASPLS*QOEARRLOSRTRERDTWDLSREQISRROWEPGMTRYLWLALLKKP SVIAKOMTASTLS*QOEARRLOSRTRERDTWDLSREQISRROWEPGMTRYLWLALLKKP SVIAKOMTASTLS*QOEARRLOSRTRERDTWDLSREQISRROMEPGMTRYLWLALLKKP **********************************	*YCVV*GITR*ISRTQ*RDVIDQSVKVRSSR*QTYLW*QWSLMQMNARHLTWFRW*RHPY *YCVV*GITR*ISRTQ*RDVIDQSVKVRSSR*QTYLW*QWSLMQMNARHLTWFRW*RHPY *YCVV*GITR*ISRTQ*RDVIDQSVKVRSSR*QTYLW*QWSLMQMNARHLTWFRW*RHPY *YCVV*GITR*ISRTQ*RDVIDQSVKVRSSR*QTYLW*QWSLMQMNARHLTWFRW*RHPY *YCVV*GITR*ISRTQ*RDVIDQSVKVRSSR*QTYLW*QWSLMQMNARHLTWFRW*RHPY *YCVV*GITR*ISRTQ*RDVIDQSVKVRSSR*QTYLW*QWSLMQMNARHLTWFRW*RHPY *YCVV*GITR*ISRTQ*RDVIDQSVKVRSSR*QTYLW*QWSLMQMNARHLTWFRW*RHPY ************************************	-MNSNLS*GWW*WLTLA*SRSILAERNNGCI*GTGSWGTSWILSTWRIICRHLTA*FRN* -MNSNLS*GWW*WLTLV*SRSILAERNNGCI*GTGSWGTSWILSTWRIICRHLTA*FS-* -MNSNLS*GWW*WLTLA*SRSILAERNNGCI*GTGSWGTSWILSTWRIICRHLTA*FRN* -MNSNLS*GWW*WLTLV*SRSILAERNNGCI*GTGSWGTSWILSTWRIICRHLTA*FRN* CMNSNLS*GWW*WLTLV*SRSILAERNNGCI*GTGSWGTSWILSTWRIICRHLTA*FRN* -MNSNLS*GWW*WLTLA*SRSILAERNNGCI*GTGSWGTSWILSTWRIICRHLTA*FRN* -MNSNLS*GWW*WLTLA*SRSILAERNNGCI*GTGSWGTSWILSTWRIICRHLTA*FRN*	IVS-FF-MTINRQV-*N*L*R*V*V*TSVW*GNYDRNYCHA*RETVTLHANLWS LNCVVVFLNDNTV*N*L*R*V*V*TSVW*GNYDRNYCHA*RETVTLHANLWS IVS-FF-MTINRQV-*N*L*R*V*V*TSVW*GNYXRNYC
Acasv-B.pep GPd58-1sv.pep GPd58-3sv.pep GPd58-4sv.pep Misv-D.pep Pavsv-D.pep	Acasv-B.pep GPd58-1sv.pep GPd58-3sv.pep GPd58-4sv.pep Misv-D.pep Pavsv-D.pep	Acasv-B.pep GPd58-1sv.pep GPd58-3sv.pep GPd58-4sv.pep M1sv-D.pep Pavsv-D.pep	Acasv-B.pep GPd58-1sv.pep GPd58-3sv.pep GPd58-4sv.pep Misv-D.pep Pavsv-D.pep
[SEQ ID NO:170] [SEQ ID NO:172] [SEQ ID NO:174] [SEQ ID NO:176] [SEQ ID NO:178]	[SEQ ID NO:170] [SEQ ID NO:172] [SEQ ID NO:174] [SEQ ID NO:176] [SEQ ID NO:180]	[SEQ ID NO:170] [SEQ ID NO:172] [SEQ ID NO:174] [SEQ ID NO:176] [SEQ ID NO:180]	[SEQ ID NO:170] [SEQ ID NO:172] [SEQ ID NO:174] [SEQ ID NO:176] [SEQ ID NO:180]

### Figure 5

9	
Figure	

AGACTGCAGTCCCGTACGCGAACGGGATACCTGGGATTTATCAAGAGAACAGATTTCA AGACTGCAGTCCCGTACGCGCAACGGGATACCTGGGATTTATCAAGAGAACAGATTTCA AGACTGCAGTCCCGGTACGCGCGAACGGGATACCTGGGATTTATCAAGAGAACAGATTTCA AGACTGCAGTCCCGTACGCGCGAACGGGATACCTGGGATTTATCAAGAGAACAGATTTCA AGACTGCAGTCCCGTACGCGCGAACGGGATACCTGGGATTTATCAAGAGAACAGATTTCA AGACTGCAGTCCCGTACGCGCAACGGGATACCTGGGATTTATCAAGAGAACAGATTTCA AGACTGCAGTCCCGTACGCGCAACGGGATACCTGGGATTTATCAAGAGAACAGATTTCA AGACTGCAGTCCCGTACGCGCAACGGGATACCTGGGATTTATCAAGAGAACAGATTTCA	CGCAGACAGATGGAGCCCGGCATGACGCGTTATTTGTGGTTGGCCCTCTTGAAGAAACCA CGCAGACAGATGGAGCCCGGCATGACGCGTTATTTGTGGTTGGCCCTCTTGAAGAAACCA CGCAGACAGATGGAGCCCGGCATGACGCGTTATTTGTGGTTGGCCCTCTTGAAGAAACCA CGCAGACAGGTGGAGCCCGGCATGACGCGTTATTTGTGGTTGGCCCTCTTGAAGAAACCA CGCAGACAGGTGGAGCCCGGCATGACGCGTTATTTGTGGTTGGCCCTCTTGAAGAAACCA CGCAGACAGGTGGAGCCCGGCATGACGCGTTATTTGTGGTTGGCCCTCTTGAAGAAACCA CGCAGACAGGTGGAGCCCGGCATGACGCGTTATTTTGTGGTTGGCCCTCTTGAAGAAACCA CGCAGACAGATGGAGCCCGGCATGACGCGTTATTTTTTGTGCTTGCCCTCTTTGAAGAAACCA	TGATATTGCGTGGTATGAGGTATCACCCGGTAGATATCGAGAACACAGTGACGAGATGTC TGATATTGCGTGGTATGAGGTATCACCCGGTAGATATCGAGAACACAGTGACGAGATGTC TGATATTGCGTGGTATGAGGTATCACCCGGTAGATATCGAGAACACAGTGACGAGATGTC TGATATTGCGTGGTATGAGGTATCACCCGGTAGATATCGAGAACACAGTGACGAGATGTC TGATATTGCGTGGTATGAGGTATCACCCGGTAGATATCGAGAACACAGTGACGAGATGTC TGATATTGCGTGGTATGAGGTATCACCCGGTAGATATCGAGAACACAGTGACGAGATGTC TGATATTGCGTGGTATGAGGTATCACCCCGGTAGATATCGAGAACACAGTGACGAGATGTC
Acasv-B GPd58-1sv GPd58-3sv GPd58-4sv Misv-D Pavsv-D	Acasv-B GPd58-1sv GPd58-3sv GPd58-4Ev Misv-D Pavsv-D	Acaev-B GPG58-1sv; GPG58-3sv GPG58-4sv Misv-D Pavsv-D
[SEQ ID NO:169] [SEQ ID NO:171] [SEQ ID NO:173] [SEQ ID NO:175] [SEQ ID NO:177]	[SEQ ID NO:169] [SEQ ID NO:171] [SEQ ID NO:173] [SEQ ID NO:175] [SEQ ID NO:177]	[SEQ ID NO:169] [SEQ ID NO:171] [SEQ ID NO:173] [SEQ ID NO:175] [SEQ ID NO:177]
	ID NO:169] Acasv-B ID NO:171] GPG58-1sv ID NO:173] GPG58-3sv ID NO:177] Misv-D ID NO:179] Pavsv-D	ID NO:169] Acasv-B ID NO:171] GPd58-1sv ID NO:173] GPd58-4sv ID NO:177] Misv-D ID NO:177] Pavsv-D ID NO:179] Acasv-B ID NO:171] GPd58-1sv ID NO:173] GPd58-1sv ID NO:173] GPd58-4sv ID NO:177] Misv-D ID NO:177] Misv-D ID NO:177] Misv-D ID NO:179] Pavsv-D

ed
$\supset$
contin
9 8
gure
Fig

ATCGATCTGTGAAAGTGCGGTCTTCACGATGACAAACCTACTTGTGGTAGCAGTGG ATCGATCAATCTGTGAAAGTGCGGTCTTCACGATGACAAACCTACTTGTGGTAGCAGTGG ATCGATCAATCTGTGAAAGTGCGGTCTTCACGATGACAAACCTACTTGTGGTAGCAGTGG ATCGATCAATCTGTGAAAGTGCGGTCTTCACGATGACAAACCTACTTGTGGTAGCAGTGG ATCGATCAATCTGTGAAAGTGCGGTCTTCACGATGACAAACCTACTTGTGGTAGCAGTGG ATCGATCAATCTGTGAAAGTGCGGTCTTCACGATGACAAACCTACTTGTGGTAGCAGTGG ATCGATCAATCTGTGAAAGTGCGGTCTTCACGATGACAAACCTACTTGTGGTAGCAGTGG A*********************************	AGCTYGATGCAGATGAACGCGAGGCACTTGACGTGGTTCCGCTGGTGACGACATCCGTAC AGCTYGATGCAGATGAACGCGAGGCACTTGACGTGGTTCCGCTGGTGACGACGACATCCGTAC AGCTYGATGCAGATGAACGCGAGGCACTTGACGTGGTTCCGCTGGTGACGACGACATCCGTAC AGCTYGATGCAGATGAACGCGAGGCACTTGACGTGGTTCCGCTGGTGACGACATCCGTAC AGCTYGATGCAGATGAACGCGAGGCACTTGACGTGGTTCCGCTGGTGACGACATCCGTAC AGCTYGATGCAGATGAACGCGAGGCACTTGACGTGGTTCCGCTGGTGACGACATCCGTAC AGCTYGATGCAGATGAACGCGAGGCACTTGACGTGGTTCCGCTGACGACATCCGTAC AGCTYGATGCAGATGAACGCGAGGCACTTGACGTGGTTCCGCTGACGACATCCGTAC	TGAATGAACAGCAACTTGTCGTAGGGGTGGTAGTGGTTGACCCTGGCGTAGTCCCGA TGAATGAACAGCAACTTGTCGTAGGGGTGGTGGTAGTGGTTGACCCTGGTGTAGTCCCGA TGAATGAACAGCAACTTGTCGTAGGGGTGGTGGTAGTGGTTGACCCTGGCGTAGTCCCGA TGAATGAACAGCAACTTGTCGTAGGGGTGGTGGTAGTGGTTGACCCTGGTGTAGTCCCGA TGTATGAACAGCAACTTGTCGTAGGGGTGGTGGTAGTGGTTGACCCTGGTGTAGTCCCGA TGTATGAACAGCAACTTGTCGTAGGGGTGGTGGTAGTGGTTGACCCTGGTGTAGTCCCGA TGAATGAACAGCAACTTGTCGTAGGGGTGGTGGTAGTGGTTGACCCTGGTGTAGTCCCGA ** **********************************	TCAATTCTCGCGGAGAAAACAACGGATGCATCTGAGGGACGGGTTCCTGGGGGGACCACT TCAATTCTCGCGGAGAGAAACAACGGATGCATCTGAGGGACGGGTTCCTGGGGGACCAGT TCAATTCTCGCGGAGAGAAAAAACGGATGCATCTGAGGGACGGGTTCCTGGGGGACCAGT TCAATTCTCGCGGAGAGAAAAAAACGATGCATCTGAGGGACGGGTTCCTGGGGGACCAGT TCAATTCTCGCGGAGAGAAAAAAACAACGGATGCATCTGAGGGACGGGTTCCTGGGGGGACCAGT TCAATTCTCGCGGAGAGAAAAAAAAAA
Acasv-B GPG58-1sv GPG58-3sv GPG58-4sv Misv-D Pavgv-D	Acasv-B GPd58-1sv GPd58-3sv GPd58-4sv Misv-D Pavsv-D	Acasv-B GPd58-1sv GPd58-3sv GPd58-4sv Misv-D Pavsv-D	Acasv-B GPd58-1sv GPd58-3sv GPd58-4sv Misv-D Pavsv-D
[SEQ ID NO:169] [SEQ ID NO:171] [SEQ ID NO:173] [SEQ ID NO:175] [SEQ ID NO:177]	[SEQ ID NO:169] [SEQ ID NO:171] [SEQ ID NO:173] [SEQ ID NO:175] [SEQ ID NO:177] [SEQ ID NO:177]	[SEQ ID NO:169] [SEQ ID NO:171] [SEQ ID NO:173] [SEQ ID NO:175] [SEQ ID NO:177]	[SEQ ID NO:169] [SEQ ID NO:171] [SEQ ID NO:173] [SEQ ID NO:175] [SEQ ID NO:177] [SEQ ID NO:179]

		•	
IGGATCCTATCTACGTGGCGTATAATATGTAGACACCTCACTGCTTAGTTT-CGTAATTG IGGATCCTATCTACGTGGCGTATAATATGTAGACACCTCACTGCTTAATTTTCGTAATTG IGGATCCTATCTACGTGGCGTATAATATGTAGACACCCTCACTGCTTAGTTT-CGTAATTG IGGATCCTATCTACGTGGCGTATAATATGTAGACACCTCACTGCTTAGTTT-CGTAATTG IGGATCCTATCTACGTGGCGTATAATATGTAGACACCTCACTGCTTAGTTT-CGTAATTG IGGATCCTATCTACGTGGCGTATAATATGTAGACACCTCACTGCTTAGTTT-CGTAATTG IGGATCCTATCTACGTGGCGTATAATATGTAGACACCTCACTGCTTAGTTT-CGTAATTG IGGATCCTATCTACGTGGCGTATAATATGTAGACACCTCACTTAGTTT-CGTAATTG IGGATCCTATCTACGTGGCGTATAATATGTAGACACCTCACTTAGTTT-CGTAATTG	AATTGTGTCGTAGTTTTTTAAATGACAATTAATAGACAAGTTTGAAATTGACTGTAGCG AATTGTGTCGTAGTTTTTTTAAATGACAACTAATAGACA-GTTTGAAATTGACTGTAGCG AATTGTGTCGTAGTTTTTTTAAATGACAATTAATAGACAAGTTTGAAATTGACTGTAGCG AATTGTGTCGTAGTTTTTTTAAATGACAATTAATAGACAAGTTTGAAATTGACTGTAGCG AATTGTGTCGTAGTTTTTTTTAAATGACAATTAATAGACAAGTTTGAAATTGACTGTAGCG AATTGTGTCGTAGTTTTTTTTAAATGACAATTAATAGACAAGTTTGAAATTGACTGTAGCG AATTGTGTCGTAGTTTTTTTAAATGACAATTAATAGACAAGTTTGAAATTGACTGTAGCG AATTGTGTCGTAGTTTTTTTAAATGACAATTAATAGACAAGTTTGAAATTGACTGTAGCG	CTAGGITTAGGTATAAACTAGCGTTTGGTAAGGCAATTATGACAGGAACTACTGTCACGC CTAGGTTTAGGTATAAACTAGCGTTTGGTAAGGCAATTATGACAGGAATTACTGTCACGC CTAGGTTTAGGTATAAACTAGCGTTTGGTAAGGCAATTATNACAGGAACTACTGTCACGC CTAGGTTTAGGTATAAACTAGCGTTTGGTAAGGCAATTATGACAGGAATTACTGTCACGC CTAGGTTTAGGTATAAACTAGCGTTTGGTAAGGCAATTATGACAGGAATTACTGTCACGC CTAGGTTTAGGTATAAACTAGCGTTTGGTAAGGCAATTATGACAGGAATTACTGTCACGC CTAGGTTTAGGTATAAACTAGCGTTTGGTAAGGCCAATTATGACAGGAACTACTGTCACGC CTAGGTTTAGGTATAAACTAGCGTTTGGTAAGGCCAATTATGACAGGAACTACTGTCACGC	GTGACGCGAGACCGTCACTTTACACGCAAACCTGTGGTCGCC GTGACGCGAGACCGTCACTTTACACGCAAACCTGTGGTCGCC GTGACGCGAGACCGTCACTTTACACGCAAACCTGTGGTCGCC GTGACGCGAGACCGTCACTTCACACGCAAACCTGTGGTCGCC GTGACGCGAGACCGTCACTTCACACGCAAACCTGTGGTCGCC GTGACGCGAGACCGTCACTTTACACGCAAACCTGTGGTCGCC
Acasv-B GPG58-1sv GPG58-3sv GPG58-4sv Misv-D Pavsv-D	Acasv-B GPd58-1sv GPd58-3sv GPd58-4sv Misv-D Pavsv-D	Acasv-B GPG58-1sv GPG58-3sv GPG58-4sv Misv-D Pavsv-D	Acasv-B GPG58-1sv GPG58-3sv GPG58-4sv M1sv-D Pavsv-D
EQ ID NO:169] EQ ID NO:171] EQ ID NO:173] EQ ID NO:175] EQ ID NO:177]	EQ ID NO:169] EQ ID NO:171] EQ ID NO:173] SQ ID NO:175] EQ ID NO:177]	2Q ID NO:169] 2Q ID NO:171] 2Q ID NO:173] 2Q ID NO:175] 3Q ID NO:177]	(Q ID NO:169] (Q ID NO:171] (Q ID NO:173] (Q ID NO:175] (Q ID NO:177] (Q ID NO:177)
Oes] Oes] Oes] Oes] Oes] Oes]	Oas]	oas] oas] oas] oas] oas]	Cas] Cas] Cas] Cas] Cas] Cas] Cas

TCCGTTATCGCTAAACAGATGACCTACAAGGTTTATATGTCAGACACGGTCAATGG TCCGTTATCGCTAAACAGATGACGCTTCAACGTTAAGTTGACAACAGGAAGCACGACGG *******************	ACACTACTITIGAGGTTGAAGGCGATGGAAAAGGAAAGGCCTTACGAGGGGGAAGACTGGCGATTTATCAAGAGAACAGATTTCA AGACTGCAGTCCCGTACGCGCGCAACGGGATACCTGGGATTTATCAAGAGAACAGATTTCA * *** * * * * * * * * * * * * * * * *	-GCAGACGG-TAAAGCTCACTGTCACCAAGGGCGGACCTCTGCCATTTGCTTG CGCAGACAGATGGAGCCCGGCATGACGCGTTATTTGTGGTTGGCCCTCTTGAAGAAACCA *****	GGATATTCTATCACACACAGAGTCAGTACGGAAGCATACCATTCACCAAGTACCCTGAAGA TGATATTGCGTGGTATGAGGTATCACCGGTAG-ATATCGAGAACACAGTGACGAGATGT ****** * * ** * ** ** ** ** ** ** ** **	CATCCCTGACTATGTAAAGCAGTCATTCCCTGAGGGATATACATGGGAGAGGATCATG CATCGATCAATCTGTGAAAGTGCGGTCTTCACGATGACAAACCTACTTGTGGTAGCAGTG **** * * * * * * * * * * * * * * * * *	AACTTCGAAGATGGTGTGTGTACTGTCAGCAATGATTCCAGCATCCAAGGTAACTGT GAGCTTGATGCAGATG-AACGCGAGGCACTTGACGTGGTTCCGCTGGTGACGACATCCGT  * * * * * * * * * * * * * * * * * * *	TTCATCTACAATGTCTCTGGTTTGAACTTTCCTCCCAATGGACC ACTGAATG-AACAGCAACTTGTCGTAGGGGTGGTGGTAGTGGTTGACCCTGGCGTAGTCC * ** ** ** ** ** ** ** ** ** ** ** ** *	TGTTATGCAAAAGACACAGGGCTGGGAACCCAACACTGGGGGTCTCTTTGCACGA CGATCAATTCTCGCGGAGAAAAAAAAACGAGATGCATCTGAGGGACGGGTTCCTGGGGGACC
Aasv-1	Aasv-1	Aasv-1	Aasv-1	Aasv-1	Aasv-1	Aasv-1	Aasv-1
Acasv-B	Acasv-B	Acaev-B	Acasv-B	Acasv-B	Acasv-B	Acasv-B	Acasv-B
[SEQ ID NO:19]	ID NO:19]	[SEQ ID NO:19]	ID NO:19]	ID NO:19]	ID NO:19]	ID NO:19]	ID NO:19]
[SEQ ID NO:169]	ID NO:169]	[SEQ ID NO:169]	ID NO:169]	ID NO:169]	ID NO:169]	ID NO:169]	ID NO:169]
ces]	čes]	(SEQ	čes]	čes] čes]	Ces]	čas]	Čas]

Figure 7

þ
nue
tin
ono
00
e 7
ure
500
H

ataggaarcaac .Tctacgtggcgt	TIGIGIGIGAAITCAAAICITACAAGGCAAAGAAGCCTGIGATGATGCCAGGGTAT TGAAITGIGICGIAGTITITITAAAIGACAATITAAIAGACAAGITITGAAAITGACAAGITGIGAAAT	CACTATGTTGAC-CGCAAGTTGGATGTAACCAATCACAACAAGGATTACACTTCC CGCTAGGTTTAGGTATAAACTAGCGTTTGGTAAGGCAATTATGACAGGAACTACTGTCAC * *** *** * * * * * * * * * * * * * *	GTTGAGCAGTGTGAAATTTCCATTGCACGCAAACCTGTGGTCGCC GC-GTGACGCGAGACCGTCACTTTACACGCAAACCTGTGGTCGCC
Aasv-1	Aasv-1	Aasv-1	Aasv-1
Acasv-B	Acasv-B	Acasv-B	Acasv-B
[SEQ ID NO:19]	[SEQ ID NO:19]	[SEQ ID NO:19]	[SEQ ID NO:19]
[SEQ ID NO:169]		[SEQ ID NO:169]	[SEQ ID NO:169]

Figure 8

SVIAKOMIYKVYMSGTVNGHYFEVEGDGKGLPYEGGQTVRLAVTKGGPLPFAWDILSPQC SVIAKQMTYKVYMSDTVNGHYFEVEGDGKGKPYEGEQTVKLIVTKGGPLPFAWDILSPQS SVIAKOMIYKVYMSGTVNGHYFEVEGDGKGKPYEGEQTVRLTVTKGGPLPFAWDILSPQS SVIAKQMIYKVYMSGTVNGHYFEVEGDGKGKPYEGEQTVKLIVTKGGPLPFAMDILSPQ\$ SVIAKQMIYKVÝMSGTVNGHYFEVEGDGKGKPYEGGEQTVKLTVTKGGPLPFAWDILSPRC SVIAKOMTYKVYMSGTVNGHYFEVEGDGKGKPYEGEQTVRLTVTKGGFLPFAWDILSPQS SVIAKQMIYKVYMSGTVNGHYFEVEGDGKGKPYEGEQTVKD.TVTKGGPLPFAWDILSPQC SVIAKQMIYKVYMSGTVNGHYFEVEGDGKGKPYEGEQTVRLTVTKGGPLPFAWDILSPQS SVIAKQMTYKVYMSGTVNGHYFEVEGDGKGKPYEGEQTVKLTVTKGGPLPFAWDIXSPQS SVIAKOMTYKVYMSGTVNGHYFEVEGDGKGKPYEGEQTVKLTVTKGGPLPFAWDILSPQC SVIAKOMIYKVYMSGTVNGHYFEVEGDGKGKPYEGEQTVKLTVIKGGPLFFAWDILSPQC SVIAKQMIYKVYMSGTVNGHYFEVEGDGKGKPYEGEQTVKLTVTKGGPLPFAWDILSPQS SVIAKQMIYXXXXSGTVXGHYFEVEGDGKGKPYEGEQTVKLTVTKGGPLPFAWDIISPQC SVIAKQMTYKVYMSGTVNGHYFEVEGDGKGKPYEGEQTVKLTVTKGGPLPFAWDILSPQS SVIAKQMTYKVYMSGTVNGHYFEVEGDGKGKPYEGEQTVRJTVTKGGPLPFAWDILSPQS SVIAKQMTYKVYMSGTVNGHYFEVEGDGKGKPYEGEQTVKLTVTEGGPLPFAWDILSPQS SVIAKQMTYKVXMSGTVNGHYFEVEGDGKGKPYEGEQTVKLTVTKGGPLPFAWDILSPQS SVIAKOMTYKVYMSGTVNGHYFEVEGDGKGKPYBGEQTVRLIVIKGGPLPFAWDILSPQS SVIAKOMIYKVYMSGTVNGHYFEVEGDGKGKPYEGEQTVRLAVTKGGPLPFAWDILSPQC SVIAKQMTYKVYMSGTVNGHYFEVEGDGKGKPYEGEQTVRLTVTKGGPLPFAWDILSPQS SVIAKQMTYKVYMSDTVNGHYFEVEGDGKGKPYEGEQTVKLTVTKGGPLPFAWDILSPQS SVIAKQMIYKVYMSGTVNGHYFEVEGDGKGRPYEGEQTVKLTVTKGGPLPFAWDILSPQS SVIAKQMIYKVYMSGTVNGHYFXVEGDGKGKPYEGEQTVKLTVTKGGPLPFAWDILSPQS SVIAKOMIYKVYMSGIVNGHYFEVQGDGKGKPYEGEQIVKLIVIKGGPLPFAMDILSPÖS SVIAKQMTYKVXMSGTVNGHYFEVEGDRKGKPYEGEQTVKLTVIKGGPLPFAWDILSPQC SVIAKQMIYKVYMSGTVNGHYFEVEGDGKGKPYEGEQTVKLTVTKGGPLPFAWDILSPQC SVIAKQMIYKVYMSGTVNGHYFEVEGDGKGKPYEGEQTVKLTVTKGGPLPFAWDILSPQC SVIAKOMIYKVYMSGTVNGHYFEVEGDGKGKPYEGEQTVKLfVTKGGPLPFAWDILSPRC SVIAKQMTYKVYMSGTVNGHYFEVEGDGKGKPYEGEQTVKLIVTKGGPLPFAWDILSPQS SVIAKQMTYKVYMSGTVNGHYFEVEGDGKGKPYEGBQTVKLTVIKGGPLPFAMDILSPQS SVIAKQMTYKVNMSGTVNGHYFEVEGDGKGKPYEGGEQTVKLTVTEGGPLPFAWDILSPQS SVIAKQMIYKVYMSGTVNGHYFEVEGDGKGKPYEGGEQTVKLIVIKGGPLPFAWDILSPQC SVIAKQMIYKVYMSGTVNGHYFEVEGDGKGKPYEGEQTVKLIVTEGGPLPFAMDILSPQS SVIAKQMTYKVYMSGTVNGHYFBAEGDGKGKPYEGEQTVKLTVTKGGPLPFAMDILSPQS SVIAKOMTYKVYMSGTVNGHYFEVEGDGKGKPYEGGQTVKLTVTKGGPLPFAMDILSPOS SVIAKQMIYKVYMSGTVNGHYFEVEGDGKGKPYEGEQTVKLIVTKGGPLPFAWDILSPQS PM1Asv-rep.pep PM1Csv-rep.pep GPd58-2sv.pep Ce61-3sv.pep Ce61-5sv.pep Ce61-7sv.pep Ce61-4sv.pep Acasv-C.pep Acasv-D.pep Acasv-A.pep LGASV-A.pep GASV-C.pep LGAsv-D.pep LGASV-E.pep Pavev-A.pep Pavsv-B.pep avev-C.pep Aasv-3.pep Aasv-P.pep Aasv-I.pep Misv-A.pep Misv-B.pep Misv-F.pep PMsv-4.pep PMsv-5.pep PPsv-1.pep PPsv-2.pep PPsv-3.pep PPsv-4.pep PPsv-5.pep PPsv-6.pep RIsv-1.pep RTsv-2.pep RIEV-3.pep Aapat2 Aapat1 NO:20] NO:26] NO:28 NO:32] NO:24 NO:30 NO:34 NO:38 NO:40 NO:44 NO:56 NO:58] NO: 66 NO: 68 NO:36 NO:42 NO:46 NO:48 NO:50 NO:54 NO: 60 NO: 62 NO: 64 NO:70 NO:72 NO: 74 NO:76 NO:78 NO:80 NO: 52 A B H A A H H A A A A A A H A A SEQ SEO SEO SEO SEO SEQ SEQ SEQ SEQ SEQ SEO SEQ SEQ SEQ SEO SEQ SEQ SEQ SEQ SEQ SEQ SEQ SEO SEQ SEQ SEQ SEO SEO SEO

### Figure 8 continued

QYGSIPFTKYPEDIPDYVKQSFPEGYTWERIMNFEDGAVCTVSNDSSIQGNCFIYHVKFS **QYGSIPFTKXPEDIPDYVKQSFPEGYTWERIVNFEDGAVCTVSNDSSIQGNCFIYNVKFS** QYGSIPFIKYPEDIPDYVKQSFPGKYTWERIMNFEDGAVCTVSNDSSIQGNCFIYHVKRS QYGSIPFIKYPEDIPDYVKQSFPEGYTWERIMNFEDGAVCTVSNDSSIQGNCFIYNVKFS QYGN IPFIKYPEDVPDYVKQSFPEGFTWERIMNFEDGAVCTVSNDSSIQGNCFIYHVKFS QYGSIPFTKYPEDIPDYVKQSFPEGYTWERIMNFEDGAVCTVSNDSSIQGNCFIYHVKFS QYGSIPFTKYPEDIPDYVKQSFPEGFTWERIMNFEDGAVCTVSNDSSIQGNCFTYHVKFS QYGSIPFIKYPEDIPDYVKQSFPEGYTWERIMNFEDGAVCTVSNDSSIQGNCFIYHVKFS QYGSXPFTKYPEDIPDYVKQSFPEGYTWERIMNFEDGAVCTVSNDSSIQGNCFIYHVKFS QYGSIPFTKXPEDIPDXVKQSFPEGFTWERIMNFEDGAVCTVSNDSSIQGNCFTYHVKFS QYGSIPFTKYPEDIPDYVKQSFPEGYTWERIMNFEDGAVCTVSNDSSIQGNCFIXNVKFS OYGSIPFIKYPEDIPDYVKQSFPGRYTWERIMNFEDGAVCTVSNDSSIQGNCFIYHVKFS QYGSIPFIKYPEDIPDYVKQSFPEGFTWERIMNFEDGAVCTVSNDSSIQGNCFIYHVKFS QYGSIPFTKYPEDIPDYVKQSFPEGYTWERIMNFEDGAVCTVSNDSSIQGNCFIYNVKFS QYGSIPFTKYPEDIPDYVKQSFPEGFTWERIMNFEDGAVCTVSNDSSIQGNCFTYHVKFS QYGSIPFTKYPEDIPDYVKQSFPEGYTWERIMNFEDGAVCTVSNDSSIQGNCFIXNVKFS QYGSIPFIKYPEDIPDYVKQSFPEGYTWBRIMNFEDGAVCTVSNDSSIQGNCFIYNVKFS QYGSIPFTKYPEDIPDYVKQSFPEGYTWERIMNFEDGAVCTVSNDSSIQGNCFIYHVKFS QYGSIPFTKYPEDIPDYVKQSFPEGYTWERIMNFEDGAVCTVSNDSSIQGNCFIYHVKFS QYGSIPFTKYPEDIPDYVKQSFPEGYTWERIMNFEDGAVCTVSNDSSIQGNCFIYHVKFS QYGSIPFTKYPEDIPDYVKQSFPEGYTWERIMNFEDGAVCTVSNDSSIQGNCFIYNVKFS QYGSIPFTKYPEDIPDYVKQSFPEGYTWERIMKFEDGAVCTVSNDSSMQGNCFIYNVKFS QYGSIPFTKYPEDIPDYVKQSFPEGYTWERIVNFEDGAVCTVSNDSSIQGNCFIYNVKFS QYGSIPFTKYPEDIPDYVKQSFPEGYTWERIMNFKDGAVCTVSNDSSIQGNCFIXNVKFS QYGNIPFTKYPEDVPDYVKQSFPEGFTWERIMNFEDGAVCTVSNDSSIQGNCFTYHVKFS QYGNIPFTKYPEDVPDYVKQSFPEGFTWERIMNFEDGAVCTVSNDSSIQGNCFTYHVKFS QYGNIPFTKYPEDVPDYVKQSFPEGFTWERIMNFEDGAVCTVSNDSSIQGNCFTYHVKFS QYGNIPFIKYPEDVPDYVKQSFPEGFTWERIMNFEDGAVCTVSNDSSIQGNCFTYHVKFS QYGSIPFTKYPEDIPDYVKQSFPEGYTWERIMNFEDGAVCTVSNDSSIQGNCFIYNVKFS QYGSIPFIKYPEDIPDYVKQSFPEGYTWERIMNFEDGAVCTVSNDSSIQGSCFIYNVKFS QYGSVPFTKYPEDIPDYVKQSFPEGYTWERIMNFEDGAVCTVSNDSSIQGNCFIYNVKFS QYGSIPFTKYPEDIPDYVKQSFPEGFTWERIMNFEDGAVCTVSNDSSIQGNCFTYHVKFS QYGSVPFTKYPEDIPDYVKQSFPEGYTWERIMNFEDGAVCTVSNDSSIQGNCFIYHVKFS <u>QYGSIPFTKYPEDIPDYVKQSFPEGYTWERIMNFEDGAVCTVSNDSSIQGNCFIYNVKFS</u> QYGSIPFTKYPEDIPDYVKQSFPEGYTWERIMNPEDGAVCTVSNDSSIQGNCFIYNVKFS oygsipftkypedipdyvkosfpegytwerimnfedgavctvsndssiogncfixnvkfs \*\* \* \*\* \*\* \*\* \*\*\*\*\*\*\*\*\* \* . \*\*\*\*\* PM1Asv-rep.pep PMICSV-rep.pep GPd58-2sv.pep Ce61-3sv.pep Ce61-5sv.pep Ce61-7sv.pep Ce61-4sv.pep Acasv-A.pep Acasv-D.pep LGASV-C.pep Acasv-C.pep LGASV-A.pep LGASV-D.pep LGASV-E.pep Pavsv-A.pep Pavsv-B.pep Pavsv-C.pep Aagv-3.pep Aagv-P.pep Aasv-1.pep Misv-A.pep Misv-B.pep Misv-F.pep PMsv-4.pep PMsv-5.pep PPsv-1.pep PPsv-2.pep PPsv-3.pep PPsv-4.pep RTSV-2.pep PPsv-5.pep PPsv-6.pep RTSV-1.pep Aapat2 Aapatı

Aasv-1.pep Aasv-3.pep

Aapat1 Aapat2

PM1Asv-rep.pep PM1Csv-rep.pep GPd58-2sv.pep Ce61-3av.pep Ce61-4sv.pep Ce61-5sv.pep Ce61-7sv.pep Acasv-D.pep Acasv-A.pep Acasv-C.pep LGABV-A.pep LGASV-C.pep LGABV-D.pep LGASV-E.pep Pavsv-A.pep Pavsv-B.pep Pavsv-C.pep Aasv-P.pep Misv-A.pep Misv-B.pep Misv-F.pep PPsv-2.pep RTsv-1.pep PMsv-4.pep PMsv-5.pep PPsv-1.pep PPsv-3.pep PPsv-4.pep PPev-5.pep PPsv-6.pep RTsv-2.pep RIEV-3.pep

glnfppngpvmokktogswepnterlfardgmlignnfmalkleggghylcefkstykark glinfppngpvmokktogfwbpnterlfardgmlignnfmalkleggghylcbfkstykakk glinfppngymgykktggwepntgrlfardgmlignnfmalkleggghylcefkstykakk glinfppngpvmqxcyqgwbpntbrllardgwlignnfwalkleggghylcefkstykark glnfppngpymokktrgwedhserlfargmlignnfmalkleggghylcgfkttykakk glnfppngpVmqkktqggwepnterlsardgmlignnfmalkleggghylcefkstykark glnfppngpvmokktoggwedhsbrlfarggmlignnfmapkleggghylcefktitykakk ginfppngpvmokktogwepnterlsardgmlignnfmalkleggghylcefkstykark glnfppngpymokktoggwedhsbrlfardgwllgnnfwalkkegggxylcbfkstykakk glinfppngpvmokktoggwedhserlfarggmlignnfmalkleggghylcefktitykakk glnfppngpvmokktoggwephserlfarggwlignnfmalkleggghylcefktitykakk glnfppngpvmokktogwepnterlfardgmlignnfwalkleggghylcefkstykakk glnfppngpvmokktggwepnterlfardgmlignnfmalkleggghylcefkstykakk gldfppngpvmokktoggwedhserlfarggwlignnfmalkleggghylcefkttykakk glnfppngpvmokktogwepnterlfardgmlignnfmalkleggghylcefkstykakk glinf ppingpumokktiogwepinterl fardgml i ginnfmalklisggghyl cefkstykakk glnfppngpvmokktoggwbpntbrlfardgwlignnfmalkleggghylcefkstykark glnfppngpvmokkkiogwepnterlfardgmlignnfmalkleggghylcefkstykark glinfppngpvmokktoggwepnterlfardgmlignnfmalkleggghylcefkstykark glnfppngpvmokktoggwepnterlyardgmlignnfmalkleggorsl--cogitmlta glnfppngpvmokktogwepnterlfardgvlignnfmalkleggghylcefkstykakk glnfppngpvmokktogwepntgrlfardgmi ignnfmalkleggghylcefkstykakk glnfppngfvmokktogwednterlyardgmlignnfmalklegsghytcefkstykakk glinf pengpvmokktogwephserlfarggml ignnfmalkleggghylcgfkttykakk glinfppngpvmokktiogfwbphserlfarggwlignnfwalkleggghylcgfktitykakk glinf ppingpumokacioggwedhserlfarggml i gnnfwalkleggghylcgfkijtykakk glnfppngpvmokktogwednterlfardgmlignnfmalkleggghylcefkstykakk glnf ppngpvmokktiogwepnterlfardgmliknnfhalkleggghylcefkstykakk glnfppngpvmokktoggwvpnterlfardgmlignnfmalkleggghylcefkstykalk glnfppngpvmokktrgwephserlfarggmlignnfmalkleggghylcgfkttykkakk glnfppngpvmokktiogwedhserlfarggmlignnfmapkleggghylcefktiykakk glnfppngpvmokktogwepnterlsardgmlignnfmalkleggghylcefkstykark glnfppngpvmokktogwbpntbrlfardgmlignnfmalklbgghylcefkstykakk ginfppngpvmokktogwepnterlfardgmlignnfmalkleggghylcefkstykakk glinf ppngpvmokktiogwednterlfardgmlignnfmalkleggghylcefkstykakk glinfpengevmokktogswepnierlyardgmlignnfmalklegggrsl\*----

### Figure 8 continued

### Figure 8 continued

Aapati	PVIOIPGYHYVDRICLDVINHNIXDYISVEOREISIARREIVACCFFRVKSRHK
Aapat2	PVKMPGYHYVDRKLDVTNHNLDYTSVEQCEISIARKPVVACRFFRVKSRHKYAVA
Aasv-1.pep	PVMMPGYHYVDRKLDVTNHNKOYTSVEQCEISIARKPVVA
Aasv-3.pep	PVKWPGYHYVDRKLDVTWHNKOYTSVEOREISIARKPVVA
Aasv-P.pep	
Acasv-A.pep	PVKMPGYHYVDRKLDVTNHNKOYISVEOCEISIARROVVA
Acasv-C.pep	PVKMPGYHCVDRICLDVTNHNKOYTSVEOREISIARKPVVA
Acasv-D.pep	PVKWPGYHYVDRKLDVTNHNKOYTSVEOCEISIARKPVVA
Ce61-3sv.pep	PVKMPGYHYVDRKLDVTNHNKOYTSVEQREISIARKPVVA
Ce61-4sv.pep	PVMMPGYHYVDRKLDVTNHNKOYTSVEQCEISIARKPVVA
Ce61-5sv.pep	PVKMPGYHYVYSTIHVTMHNKDYTSVEQCEISXXRKPVVA
Cecl-7sv.pep	PUKAPGYHYVDRKLDVINHAKOYISVEQCEISIARKPVVA
GPd58-2sv.pep	PVMMPGYHYVDRKLDVTNHNKDYTSVEQCEISIARKPVVA
LGASV-A.pep	PVMMPGYHYVDRICLDVINENKDYTSVEQCEISIARICPVVA
LGASV-C.pep	PVKMPGYHYVDRKIDVTNHNKDYTSVEQCEISIARKPVVR
LGASV-D.pep	PVMMPGYHYVDRKLDVINENKDYTSVEQCEISIARKPVVA
LGASV-E.pep	PVMMPGYHYVDRKLDVTNHNKOYTSVEQCEISIARKPVVA
Misv-A.pep	PVKMPGYHYVDRKLDVINHNKOYISVEQREISIARKPVVA
Misv-B.pep	PVKMPGYHYVDRKLDVTNHNKOYTSVEQREISIARKPVVA
Misv-F.pep	PVKMPGYHYVDRKLDVTMHNKOYTSVEQREISIARKPVVA
PMIAsv-rep.pep	PVMMPGYHYVDRKLDVTNHNKOYTSVEQCEISIARKPVVA
PM1Csv-rep.pep	PVMMPGYHYVDRKLDVTNHNKDYTSVEQCEISIARKPVVA
PMsv-4.pep	NWM-PITTRITLPLSSVRFPSHANLMSDDVFSESNQGT
PMsv-5.pep	PVMMPGYHYVDRKLDVTNHNKOYTSVEQCEISIARKPVVA-RFFRVKSRHK
PPsv-1.pep	PVKMPGYHYVDRKLDVTNHNKOYISVEQCEISIARKPVVA
PPsv-2.pep	PVKMPGYHYVDRKCLDVINHNKOYISVEQCEISIARKPVVA
PPsv-3.pep	PVKMPGYHYVDRKLDVTNHNKDYISVEQCETSIARKPVVA
PPsv-4.pep	PVKMPGYHYVDRKLDVINHNKOYISVEQCEISIARKPVVA
PPsv-5.pep	PVMMPGYHYVDRKCLDVINHNKOYTSVEQCEISIARKPVVA
PPsv-6.pep	PVMMPGYHYVDRKLDVTINHNKDYTSVEQCGISIARKPVVA
Pavev-A.pep	PVMMPGYHYVDRKLDVTNHNKOYTSVEQCEISIARKPVVA
Pavsv-B.pep	PVKMPGYHYVDRKLDVINHNKOYISVEQCEISIARKPVVA
Pavsv-C.pep	PVKMPGYHYVDRKLDVTNHNKDYTSVEQREISIARKPVVA
RTsv-1.pep	PVMMPGYHYVDRKCLDVTNHNKDYTSVEQCEIPIARKPVVA
RTsv-2.pep	PVMMPGYHYVDRKLDVINHNKOYTSVEQCEISIARKÞVVA
RTsv-3.pep	PVMMPGYHYVDRKCLDVTNHNKOYTSVEQCEISIARKPVVA
	-

SEQ ID NO	N-terminus	name	extra bases	stop codon at AA14
SEQ ID NO:170	svlak	Acasv-B.pep	701 bases	SVIAKQMTASTLS*
SEQ ID NO:171	sviak	GPd58-1sv.pep	701 bases	SVIAKOMTASTLS.
SEQ ID NO:172	sviak	GPd58-3sv.pep	701 bases	SVIAKOMTASTLS*
SEQ ID NO:173	sviak	GPd58-4sv.pep	701 bases	SVIAKOMTASPLS*
SEQ 10 NO:174	sviak	Misv-D.pep	701 bases	SVIAKOMTASTLS*
SEQ ID NO:180	svíak	Pavsv-D.pep	701 bases	SVIAKOMTASTLS*
	N-terminus	name	type	"QYG" fluorophore and 26 as within 5A
SEQ ID NO:64	svíak	PPsv-1.pep	5	QVLSPQCQYGNIFWRNSYEHENMGRLQCE
SEQ ID NO:65	sviak	PPsv-2.pep	ស	QVLSPQCQYGNIFWRNSYEHENMGRLQCE
SEQ ID NO:66	sviak	PPsv-3.pep	ស	QVLSPQCQYGNIFWRNSYEHENMGRLQCE
SEQ ID NO:67	sviak	Acasv-A.pep	4	QVLSPRCQYGNIFWRNSYEHENMGRLQCE
SEQ ID NO:68	sviak	PPsv-4.pep	4	OVLSPRCQYGNIFWRNSYEHENMGRLOCE
SEQ ID NO:69	sviak	Acasv-D.pep	Ψ-	OVLSPOCOYGSIFWRNSYEHENMERLOCE
SEQ ID NO:70	sviak	LGAsv-C.pep	<b>-</b> -	QVLSPQCQYGSIFWRNSYEHENMERLOCE
SEQ ID NO:71	sviak	Pavsv-B.pep	~	QVLSPQCQYGSIFWRNSYEHENMERLOCE
SEQ ID NO:72	sviak	Ce61-7sv-rep.pel	RES 17	QVLSPQCQYGSIFWRNSYEHENMERLQCE
SEQ ID NO:73	sviak	Ce61-5sv-rep.per	RES 18	QVLSPQCQYGSIFWRNSYEHENMESIQCE
SEQ ID NO:74	sviak	Ce61-4sv.pep	20	QVXSPQSQYGSXYWRNSYEHENMERLOCE
SEQ ID NO:75	sviak	PMsv-4.pep	<b>5</b> *	QVLSPQSQYGSIYWRNSYENENM*
SEQ ID NO:76	sviak	LGAsv-E.pep	7	QVLSPQSQYGSIYWRNSYENENMERLOCE
SEQ ID NO:77	svíak	RTsv-1.pep	7	QVLSPQSQYGSIYWRNSYENENMERLOCE
SEQ ID NO:78	sviak	GPd58-2sv.pep	7	QVLSPQSQYGSIYWRNSYENENMERLOCE
SEQ ID NO:79	sviak	LGAsv-A.pep	. 73	QVLSPQSQYGSIYWRNSYENENMERLQCE
SEQ ID NO:80	sviak	LGAsv-D.pep	7	QVLSPQSQYGSIYWRNSYENENMERLQCE
SEQ ID NO:81	sviak	PPsv-5.pep	7	QVLSPQSQYGSIYWRNSYENENMERLQCE
SEQ ID NO:82	sviak	RTsv-2.pep	7	QVLSPQSQYGSIYWRNSYENENMERLQCE
SEO ID NO:83	sviak	RTsv-3.pep	7	QVLSPQSQYGSIYWRNSYENENMERLQCE
SEQ ID NO:84	sviak	Aasv-P.pep	74	QVLSPQSQYGSIYWRNSYENENMERLQCE
SEQ ID NO:85	sviak	PMsv-5.pep	7	QVLSPQSQYGSIYWRNSYENENMERLQCE
SEC ID NO:86	svlak	PM1Csv-rep.pep	7	<b>QVLSPQSQYGSPYWRNSYENENMERLQCE</b>
SEQ ID NO:87	sviak	Aasv-1.pep	15.	QVLSPQSQYGSIYWRNSYENGNMERLQCE
SEQ ID NO:88	sviak	PM1Asv-rep.pep	15	QVLSPQSQYGSIYWRNSYENGNMERLQCE
SEQ ID NO:89	sviak	PPsv-6.pep	9	QVLSPQSQYGSIYWRNSYENENMERLQCG
SEQ ID NO:90	sviak	Misv-A.pep	14	QVLSPQSQYGSIYWRNSYENENMERLQRE
SEQ ID NO:91	sviak	Aasv-3.pep	14	QVLSPQSQYGSIYWRNSYENENMERLQRE
SEQ ID NO:92	sviak	Acasv-C.pep	4	<b>QVLSPQSQYGSIYWRNSYENENMERLQRE</b>
SEQ ID NO:93	sviak	Ce61-3sv.pep	44	QVLSPQSQYGSIYWRNSYENENMERLQRE
SEQ ID NO:94	sviak	Misv-B.pep	4	QVLSPQSQYGSIYWRNSYENENMERLQRE
SEO ID NO:95	sviak	Misv-F.pep	4	QVLSPQSQYGSIYWRNSYENENMERLQRE
SEQ ID NO:96	sviak	Pavsv-C.pep	17	QVLSPQSQYGSVYWRNSYENENMERLQRE
SEQ ID NO:97	sviak	Pavsv-A.pep	7	<b>QVLSPQSQYGSVYWRNSYVNENMERLQCE</b>

### Figure 9

Variable ar VIAKMKVYMSGVNFEVE GKKYQVKTKWLSPOCOYGNIEVDK/ EWIEDIA	1 1
SEQ ID NO:65	GKKYQVKTKWLSPQQQYGNIEVDK/FWERMNEDISNDSINTYHNMQQE HSEFGGMLINMLLEGGLGTKKPVKYVDRKT
SEQ 10 NO:66	GKKYQVKTKWLSPQCQYGNIEVDKK FWERMNEDTSNDSINTYHNMOOF HSFFGGMI NMI I FGG GTKKDAXAXDDBK
SEC ID NO:67	
SEC ID NO:69	HSEFGGMLINMLLEGGLGTKKPVKYVDRKL
SEQ ID NO:70	GKKYOVKTAM SPOONSEIEDKO FWERMIEDTSNDSINTYHNMOOE HSEFGGMLINMPLEGGLETKKPVKYVDRKL
SEQ ID NO:71	GKKYQVKTKWLSPQCQYGSIEIDKO FWEKMNED I SNDSIN I YHDMQQE HSEFGGMLINMLLEGGLETKKPVKYVDRKL
SEQ ID NO:72	GKKYQVKTKWLSPQCQYGSIEDKQ FWERMNEDTSNDSIN FINMAGE HSEFGGM. INM. I ECCL FTAZA 1901-1901
SEQ ID NO:73	GKKYQVKTKWLSPQCQYGSIEIDKQ FWERMNEDTSNDSINTYHNMOOF HSFEGGM INMILLEGGLETKKYAKVOKKT
SECTIONO:74	GKKYQVKTXWXSPQSQYGSXEIDKC YWERMNEDTSNDSINIYHNMQQE
SEQ ID NO:76	NTEYDGMLINMLLEGG?
SEO ID NO:77	GANTAYA EWLAPUSAY GSIEIDKA YWERMNEOTSNOSINIYANMADE
SEQ ID NO:78	GKKYON/HAWA GOODSOON SHOWNED INDIVIDUE
SEQ ID NO:79	GRANDWICH SOUND SOUND SOUND SWERMING TO SUNIVINIMODE
SEQ ID NO:80	GKKYOVITKMI SEDSONOSITEDINO VITTURE WEBSINIYANIMQOE NTEFDGMLINMLLEGGLESKKPVMYVDRKL
SEQ ID NO:81	GKKYOVKTKWI SPOSOVOSIEJOKO WIESELISTOKOSIANIKO GKKYOVKTKWI SPOSOVOSIEJOKO WIESELISTOKOSIANIKO SPOSOVOSIEJOKO WIESELISTOKOSIANIKO SPOSOVOSIANIKO SPOSOVOSIANIKA SPOSOVOSIANIKO SPOSOVOSIANIKO SPOSOVOSIANIKO SPOSOVOSIANIKA SPOSOVOSIANIKA SPOSOVOSIANI
SEQ ID NO:82	GKKYQVKTWI SPOSOVSEIEIDKO VALEBARIEDISHIVANINGOE NIEFDGAILINAILLEGGLESKKPVMYVDRA
SEQ ID NO:83	GKKYQVKTXWLSPOSOYGSIFICKO XWEBANAIGHTSINTONINACIOE NIEFDGMLINMLLEGGLESKKPVMYVDRW
SEQ ID NO:84	GKKYQVRTKWLSPOSOYSSIEIDKO XWEBMANEDTSNIPSNINGEN I FELGMILIMMILEGGLESKKPVMYVDRKI
SEQ ID NO:85	GKKYQVKTKWL3PQSQYGSIEIDKQ YWEBIANIKYTSINICATININACIE YN EFUGAILLINALLEGGLESKKYVRYVDRY. DTTEQCEISIARK
SEQ ID NO:86	GKRYQVKTKWLSPQSQYGSIEDKC YWERMKEDTSONDSANIYONINGO NIEDGYCLWILLEGGLESKFYWYDRKL
SEQ ID NO:87	GKKYQVKTKWLSPQSQYGSIEIDKQ YWERMAEDTSNDSINTNINMOODE NTGEDOM INMIT TO OF TOWN YOURK
SEQ ID NO:88	GKKYQVKTKWLSPQSQYGSIEIDKQ YWERMNEDTSNDSINKNINGOG NTGEFOON INN I TOOL TOOL TOOL TOOL TOOL TOOL TOOL TO
SEO ID NO:89	GKKYQVKTKWLSPQSQYGSIEDKQ YWERMNEDTSNDSISYNNNOCE NTEEDGALLING TOOL TOOL TOOL TOOL TOOL TOOL TOOL TOO
SEQ ID NO:90	Ÿ
	GKKYQVRTKWLSPQSQYGSIEIDKQ YWERMNEDTSNDSINIYHNMQQE NTELDGMLNMILLEGGLESKRPVXYYDRY
SECON CLUDES	GKKYQVRTKWLSPQSQYGSIEIDKQ YWERMNEDTSNDSINIYHNMODE NTESDGMI INMI I EGG EGKDBYKYYNDED
SEQ ID NO:93	GKKYQVRTKWLSPQSQYGSIEIDKQ YWERMNEDTSNDSINYHNMAODE NTESDEMI IAMII EGGI EGGESAKTYKVVOUKKL
SEQ ID NO:94	GKKYQVRTKWLSPQSQYGSIEIDKQ YWERMNEDTSNDSINIYHNMQQE
SEO ID NO:85	NTEFDGMLINMLLEGGLESKRPVKYVDRK
SEQ ID NO:97	ON TUNK! EWEST GSOSOSOS HERE TO SWERMEDTSNDSINIYHINMOOE NTESDGMLINMLEGGLESKRPVKYVDRKL
	THE THE STATE OF T

## Figure 9 continued

SEQ ID NO.92         spial         Aams-6,pep         1         QULSPOCOY GSIPWRISYS FHENMERLQCE           SEQ ID NO.93         syiat         Coms-6,pep         13         QULSPOCOY GAIIPWRISY FHENMERLQCE           SEQ ID NO.93         syiat         Cenns-6,pep         5         QULSPOCOY GAIIPWRISY FHENMERLQCE           SEQ ID NO.93         syiat         Cenns-6,pep         5         QULSPOCOY GAIIPWRISY FHENMERLQCE           SEQ ID NO.93         syiat         Cenns-6,pep         6         QULSPOCOY GAIIPWRISY FHENMERLQCE           SEQ ID NO.93         syiat         Cenns-6,pep         16         QULSPOCOY GAIIPWRISY FHENMERLQCE           SEQ ID NO.93         syiat         Cenns-6,pep         1         QULSPOCOY GSIPWRISY FHENMERLQCE           SEQ ID NO.93         syiat         Mims-2,pep         1         QULSPOCOY GSIPWRISY FHENMERLQCE           SEQ ID NO.100         syiat         Mims-2,pep         1         QULSPOCOY GSIPWRISY FHENMERLQCE           SEQ ID NO.100         syiat         RTims-6,pep         1         QULSPOCOY GSIPWRISY FHENMERLQCE           SEQ ID NO.101         syiat         Amms-6,pep         1         QULSPOCOY GSIPWRISY FHENMERLQCE           SEQ ID NO.101         syiat         Amms-6,pep         1         QULSPOCOY GSIPWRISY FHENMERLQCE	SEQ ID NO	N-terminus	name	type	"QYG" fluorophore and 26 aa within 5A
NO:93 sviat PPms-G.pep 13  NO:94 sviat Cems-G.pep 5  NO:95 sviat Cems-F.pep 5  NO:96 sviat Cems-F.pep 5  NO:98 sviat Cems-F.pep 5  NO:99 sviat PPms-E.pep 16  NO:103 sviat Mims-A.pep 17  NO:103 sviat Mims-C.pep 17  NO:104 sviat Mims-C.pep 17  NO:105 sviat Mims-C.pep 17  NO:106 sviat Aams-C.pep 17  NO:107 sviat PPd57-2.ms.pep 17  NO:108 sviat PPd57-3.pep 17  NO:109 sviat PMms-E.pep 17  NO:101 sviat PPd57-2.ms.pep 17  NO:110 sviat PMms-E.pep 2  NO:110 sviat PMms-E.pep 2  NO:111 sviat PMms-E.pep 2  NO:112 sviat PMms-C.pep 9  NO:114 sviat PMms-C.pep 19  NO:115 sviat PMms-C.pep 19  NO:116 sviat PMms-C.pep 6  NO:117 sviat PMms-C.pep 6  NO:118 sviat PMms-C.pep 6  NO:119 sviat PMms-C.pep 6  NO:110 sviat Pavms-C.pep 6  NO:111 sviat Pavms-C.pep 6  NO:112 sviat Acams-C.pep 6  NO:120 sviat Pavms-C.pep 6  NO:121 sviat Pavms-C.pep 6  NO:122 sviat Pavms-C.pep 6  NO:123 sviat Acams-C.pep 6  NO:124 sviat Pavms-C.pep 6  NO:125 sviat Acams-C.pep 6  NO:126 sviat Acams-C.pep 6  NO:127 sviat Acams-C.pep 6  NO:128 sviat Acams-S.pep 6  NO:129 sviat Acams-S.pep 6  NO:129 sviat Acams-S.pep 6  NO:129 sviat Acams-S.pep 6  NO:120 sviat Acams-S.pep 6  NO:121 sviat Acams-S.pep 6  NO:122 sviat Acams-S.pep 6  NO:123 sviat Acams-S.pep 6  NO:124 sviat Acams-S.pep 6  NO:125 sviat Acams-S.pep 6	SEQ ID NO:92	sgiat	Aams-5.pep	1	QVLSPQCQYGSIFWRNSYEHENMERLQCE
NO:94 sviat Cems-G.pep 5  NO:95 sviat Cems-H.pep 5  NO:96 sviat PPms-E.pep 5  NO:98 sviat PPms-T.pep 6  NO:99 sviat Pems-T.pep 16*  NO:100 sviat Mims-A.pep 1  NO:101 sviat Mims-A.pep 1  NO:101 sviat Mims-C.pep 1  NO:102 sviat Mims-C.pep 1  NO:103 sviat Mims-C.pep 1  NO:104 sviat Mims-C.pep 1  NO:105 sviat PPd57-2ms, pep 1  NO:106 sviat PPd57-3.pep 1  NO:107 sviat PMms-B.pep 2  NO:108 sviat PMms-C.pep 1  NO:109 sviat PMms-C.pep 1  NO:110 sviat PMms-C.pep 2  NO:111 sviat PMms-C.pep 2  NO:112 sviat PMms-C.pep 1  NO:114 sviat PMms-C.pep 1  NO:115 sviat PAmms-C.pep 6*  NO:116 sviat Pamms-C.pep 6*  NO:117 sviat Aams-C.pep 6  NO:118 sviat Acams-C.pep 6  NO:119 sviat Acams-C.pep 6  NO:110 sviat Acams-C.pep 6  NO:111 sviat Acams-C.pep 6  NO:112 sviat Acams-C.pep 6  NO:113 sviat Acams-C.pep 6  NO:114 sviat Acams-C.pep 6  NO:115 sviat Acams-C.pep 6  NO:115 sviat Acams-C.pep 6  NO:116 sviat Acams-C.pep 6  NO:117 sviat Acams-C.pep 6  NO:118 sviat Acams-C.pep 6  NO:119 sviat Acams-C.pep 6  NO:110 sviat Acams-C.pep 6  NO:111 sviat Acams-C.pep 6  NO:112 sviat Acams-C.pep 6  NO:112 sviat Acams-C.pep 6  NO:113 sviat Acams-C.pep 6  NO:112 sviat Acams-C.pep 6	SEQ ID NO:83	sviat	PPms-G.pep	13	QVLSPQCQYGNIFWGNSYEHENMGRLQCE
NO:95 sviat Cems-H.pep 5  NO:96 sviat PPms-E.pep 5  NO:98 sviat PPms-1.pep 6  NO:99 sviat Cems-1.pep 16*  NO:100 sviat Mims-A.pep 1  NO:101 sviat Mims-A.pep 1  NO:102 sviat Mims-A.pep 1  NO:103 sviat Mims-C.pep 1  NO:104 sviat Mims-C.pep 1  NO:105 sviat Aams-C.pep 1  NO:105 sviat PMms-A.pep 1  NO:106 sviat PMms-A.pep 1  NO:107 sviat PMms-B.pep 2  NO:108 sviat PMms-B.pep 2  NO:110 sviat PMms-B.pep 2  NO:111 sviat PMms-E.pep 1  NO:112 sviat PMms-C.pep 1  NO:114 sviat PMms-C.pep 1  NO:115 sviat PMms-C.pep 1  NO:116 sviat PMms-C.pep 6*  NO:117 sviat CGAms-C.pep 6*  NO:118 sviat Pavms-A.pep 6*  NO:119 sviat Acams-A.pep 6*  NO:120 sviat Acams-C.pep 6  NO:131 sviat Acams-C.pep 6  NO:132 sviat Acams-C.pep 6  NO:123 sviat Acams-C.pep 6  NO:124 sviat Acams-C.pep 6  NO:125 sviat Acams-S.pep 6  NO:125 sviat Acams-S.pep 6  NO:126 sviat Acams-S.pep 6  NO:127 sviat Acams-S.pep 6  NO:128 sviat Acams-S.pep 6  NO:129 sviat Acams-S.pep 6  NO:120 sviat Acams-S.pep 6  NO:121 sviat Acams-S.pep 6  NO:1229 sviat Acams-S.pep 6  NO:1230 sviat Acams-S.pep 6	SEQ ID NO:94	sviat	Cems-G.pep	S	QVLSPQCQYGNIFWRNSYEHENMGRLQCE
NO:96 sviat PPms-Epep 5  NO:97 sviat Cems-F.pep 5  NO:98 sviat Cems-I.pep 8  NO:100 sviat Mims-A.pep 16*  NO:101 sviat Mims-A.pep 11  NO:102 sviat Mims-C.pep 11  NO:103 sviat Mims-C.pep 11  NO:104 sviat RTms-A.pep 11  NO:105 sviat Aams-C.pep 11  NO:106 sviat Aams-C.pep 11  NO:107 sviat PMms-Epep 11  NO:108 sviat PMms-Epep 11  NO:110 sviat PMms-Epep 12  NO:110 sviat PMms-C.pep 12  NO:111 sviat PMms-C.pep 12  NO:112 sviat PMms-C.pep 13*  NO:113 sviat Pavms-C.pep 6*  NO:114 sviat Pavms-C.pep 6*  NO:115 sviat Acams-C.pep 6*  NO:115 sviat Acams-C.pep 6*  NO:116 sviat Pavms-C.pep 6*  NO:117 sviat Acams-C.pep 6*  NO:118 sviat Acams-C.pep 6*  NO:119 sviat Acams-C.pep 6*  NO:120 sviat Acams-C.pep 6*  NO:121 sviat Acams-C.pep 6*  NO:122 sviat Pavms-C.pep 6*  NO:123 sviat Acams-C.pep 6*  NO:124 sviat Pavms-C.pep 6*  NO:125 sviat Acams-C.pep 6*  NO:125 sviat Acams-C.pep 6*  NO:126 sviat Acams-C.pep 6*  NO:127 sviat Acams-C.pep 6*  NO:128 sviat Acams-C.pep 6*  NO:129 sviat Acams-C.pep 6*  NO:120 sviat Acams-C.pep 6*  NO:121 sviat Acams-C.pep 6*  NO:122 sviat Acams-C.pep 6*  NO:123 sviat Acams-C.pep 6*  NO:124 sviat Acams-C.pep 6*  NO:125 sviat Acams-C.pep 6*  NO:126 sviat Acams-C.pep 6*  NO:127 sviat Acams-C.pep 6*  NO:128 sviat Acams-C.pep 6*  NO:129 sviat Acams-C.pep 6*  NO:120 sviat Acams-C.pep 6*  NO:121 sviat Acams-C.pep 6*  NO:122 sviat Acams-C.pep 6*  NO:1230 sviat Acams-C.pep 6*  NO:124 sviat Acams-C.pep 6*  NO:125 sviat Acams-C.pep 6*  NO:126 sviat Acams-C.pep 6*  NO:127 sviat Acams-C.pep 6*  NO:128 sviat Acams-C.pep 6*  NO:129 sviat Acams-C.pep 6*  NO:120 svi	SEQ ID NO:95	sviat	Cems-H.pep	2	QVLSPQCQYGNIFWRNSYEHENMGRLQCE
NO:97         sviat         Cems-F.pep         5           NO:98         sviat         PPms-1.pep         8           NO:99         sviat         Mi68Dms.pep         1           NO:100         sviat         Mims-A.pep         1           NO:101         sviat         Mims-B.pep         1           NO:102         sviat         Mims-C.pep         1           NO:103         sviat         RTms-S.pep         1           NO:104         sviat         RTms-C.pep         1           NO:105         sviat         RTms-C.pep         1           NO:106         sviat         PMms-A.pep         1           NO:107         sviat         PMms-A.pep         1           NO:108         sviat         PMms-B.pep         2           NO:110         sviat         PMms-C.pep         1           NO:111         sviat         PMms-C.pep         1           NO:112         sviat         PMms-C.pep         1           NO:113         sviat         PMms-C.pep         6           NO:124         sviat         Pms-C.pep         6           NO:125         sviat         Pms-C.pep         6	SEQ (D NO:96	sviat	PPms-E.pep	2	<b>QVLSPQCQYGNIFWRNSYEHENMGRLQCE</b>
NO:98         sviat         PPms-1.pep         8           NO:99         sviat         Cems-1.pep         16*           NO:100         sviat         MiR8-M.pep         1           NO:101         sviat         Mims-A.pep         1           NO:102         sviat         Mims-B.pep         1           NO:103         sviat         RTms-5.pep         1           NO:104         sviat         RTms-5.pep         1           NO:105         sviat         RTms-4.pep         1           NO:106         sviat         PMs-A.pep         1           NO:107         sviat         PMs-A.pep         1           NO:108         sviat         PMs-A.pep         1           NO:109         sviat         PMms-A.pep         1           NO:110         sviat         PMms-A.pep         2           NO:111         sviat         PMms-A.pep         1           NO:112         sviat         PMms-C.pep         1           NO:115         sviat         PMms-C.pep         1           NO:116         sviat         PMms-C.pep         1           NO:120         sviat         Acams-C.pep         6	SEQ ID NO:97	sviat	Cems-F.pep	ນ	<b>QVLSPQCQYGNIFWRNSYEHENMGRLQCE</b>
NO:99         sviat         Cems-l.pep         16*           NO:100         sviat         Mi68Dms.pep         1           NO:101         sviat         Mims-A.pep         1           NO:102         sviat         Mims-B.pep         1           NO:103         sviat         Mims-C.pep         1           NO:104         sviat         RTms-S.pep         1           NO:105         sviat         RTms-S.pep         1           NO:106         sviat         PMs-B.pep         1           NO:107         sviat         PMs-B.pep         1           NO:108         sviat         PMms-B.pep         1           NO:109         sviat         PMms-B.pep         2           NO:110         sviat         PMms-B.pep         2           NO:111         sviat         PMms-B.pep         1           NO:112         sviat         PMms-C.pep         18*           NO:114         sviat         PMms-C.pep         18*           NO:115         sviat         PMms-C.pep         18*           NO:116         sviat         PMms-C.pep         6*           NO:120         sviat         Pavms-C.pep         6*	SEQ ID NO:98	sviat	PPms-1.pep	80	RVLSPQCQYGNIFWRNSYEHENMGRLQCE
NO:100         sviat         MißBDms.pep         1           NO:101         sviat         Mims-A.pep         1           NO:102         sviat         Mims-B.pep         1           NO:103         sviat         Mims-C.pep         1           NO:104         sviat         RTms-C.pep         1           NO:105         sviat         Aams-L.pep         1           NO:106         sviat         PMds7-2ms.pep         1           NO:107         sviat         PMds7-2ms.pep         1           NO:108         sviat         PMds7-2ms.pep         1           NO:109         sviat         PMds7-4ms.pep         2           NO:110         sviat         PMms-E.pep         2           NO:111         sviat         PMms-C.pep         9           NO:115         sviat         PMms-C.pep         18*           NO:116         sviat         PMms-C.pep         18*           NO:117         sviat         PMms-C.pep         6*           NO:120         sviat         Pavms-C.pep         6*           NO:121         sviat         Pavms-S.pep         6*           NO:122         sviat         Pavms-S.pep         6* </td <td>SEQ ID NO:99</td> <td>sviat</td> <td>Cems-l.pep</td> <td>16*</td> <td>QVLSPQCQYGNIFWRNSYEHENMERLQCE</td>	SEQ ID NO:99	sviat	Cems-l.pep	16*	QVLSPQCQYGNIFWRNSYEHENMERLQCE
NO:101  NO:102  Sviat Mims-B.pep 1  NO:103  Sviat Mims-C.pep 1  NO:104  Sviat RTms-C.pep 1  NO:105  Sviat Aams-C.pep 1  NO:106  Sviat Aams-C.pep 1  NO:108  Sviat Aams-C.pep 1  NO:109  Sviat PMms-B.pep 1  NO:110  Sviat PMms-B.pep 2  NO:111  Sviat PMms-B.pep 2  NO:112  Sviat PMms-C.pep 2  NO:114  Sviat PMms-C.pep 19*  NO:115  Sviat PAms-C.pep 6*  NO:116  Sviat PAms-C.pep 6*  NO:120  Sviat Acams-C.pep 6*  NO:121  Sviat Acams-C.pep 6*  NO:122  Sviat Acams-S.pep 6*  NO:124  Sviat Acams-S.pep 6*  NO:125  Sviat Acams-S.pep 6  NO:126  Sviat Acams-S.pep 6  NO:127  Sviat Acams-S.pep 6  NO:128  Sviat Acams-S.pep 6  NO:129  Sviat Acams-S.pep 6  NO:120  Sviat Acams-S.pep 6  NO:121  Sviat Acams-S.pep 6  NO:122  Sviat Acams-S.pep 6  NO:124  Sviat Acams-S.pep 6  NO:125  Sviat Acams-S.pep 6  NO:126  Sviat Acams-S.pep 6  NO:127  Sviat Acams-S.pep 6  NO:128  Sviat Acams-S.pep 6  NO:129  Sviat Acams-S.pep 6  NO:120  Sviat Acams-S.pep 6  NO:121  Sviat Acams-S.pep 6  NO:122  Sviat Acams-S.pep 6  NO:123  Sviat Acams-S.pep 6  NO:124  Sviat Acams-S.pep 6  NO:125  Sviat Acams-S.pep 6  NO:126  Sviat Acams-S.pep 6  NO:127  Sviat Acams-S.pep 6  NO:128  Sviat Acams-S.pep 6  NO:129  Sviat RTms-C.pep 6  NO:120  Sviat RTms-C.pep 6  NO:120  Sviat Acams-S.pep 6  NO:120  Sviat A	SEQ ID NO:100	sviat	Mi68Dms.pep	-	<b>QVLSPQCQYGSIFWRNSYEHENMERLQCE</b>
NO:102         svlat         Mims-B.pep         1           NO:103         sviat         Mims-C.pep         1           NO:104         sviat         RTms-S.pep         1           NO:105         sviat         Aams-L.pep         1           NO:106         sviat         Aams-G.pep         1           NO:107         sviat         PPd57-2ms.pep         1           NO:108         sviat         PPd57-2ms.pep         1           NO:109         sviat         PPMms-A.pep         2           NO:110         sviat         PPMms-A.pep         2           NO:111         sviat         PPMms-C.pep         2           NO:112         sviat         PPMms-C.pep         18*           NO:115         sviat         PPMms-C.pep         18*           NO:116         sviat         PPMms-C.pep         6*           NO:117         sviat         PPMms-C.pep         6*           NO:118         sviat         Pavms-C.pep         6*           NO:120         sviat         Pavms-S.pep         6*           NO:121         sviat         Pavms-S.pep         6           NO:122         sviat         Acams-S.pep         6<	SEQ ID NO:101	sviat	Mims-A.pep		QVLSPQCQYGSIFWRNSYEHENMERLQCE
NO:103 sviat Mims-C.pep 1  NO:104 sviat RTms-5.pep 1  NO:105 sviat Aams-2.pep 1  NO:106 sviat Aams-4.pep 1  NO:107 sviat Aams-6.pep 1  NO:108 sviat PPd57-2ms.pep 1  NO:109 sviat PPMms-B.pep 2  NO:110 sviat PMms-B.pep 2  NO:111 sviat PMms-B.pep 2  NO:112 sviat PPMms-C.pep 9  NO:114 sviat PPMms-C.pep 9  NO:115 sviat PPms-2.pep 19*  NO:116 sviat PPms-2.pep 6*  NO:120 sviat Acams-4.pep 6*  NO:121 sviat Acams-3.pep 6  NO:122 sviat RTms-1.pep 6  NO:124 sviat RTms-1.pep 6  NO:125 sviat Acams-3.pep 6  NO:126 sviat RTms-1.pep 6  NO:127 sviat Acams-5.pep 6  NO:128 sviat Acams-5.pep 6  NO:129 sviat Acams-5.pep 6  NO:120 sviat Acams-5.pep 6  NO:121 sviat Acams-5.pep 6  NO:122 sviat RTms-1.pep 6  NO:123 sviat Acams-5.pep 6  NO:124 sviat Acams-5.pep 6  NO:125 sviat Acams-5.pep 6	SEQ ID NO:102	sviat	Mims-B.pep	τ-	<b>QVLSPQCQYGSIFWRNSYEHENMERLQCE</b>
NO:104         sviat         RTms-5.pep         1           NO:105         sviat         Aams-2.pep         1           NO:106         sviat         Aams-4.pep         1           NO:107         sviat         Aams-6.pep         1           NO:108         sviat         PPd57-2ms.pep         1           NO:109         sviat         PMms-B.pep         2           NO:110         sviat         PMms-B.pep         2           NO:111         sviat         PMms-D.pep         2           NO:112         sviat         PMms-D.pep         2           NO:114         sviat         PMms-C.pep         12           NO:115         sviat         PMms-C.pep         9           NO:116         sviat         PMms-C.pep         14           NO:117         sviat         PRAms-C.pep         6*           NO:120         sviat         Pavms-A.pep         11           NO:121         sviat         Pavms-A.pep         6*           NO:122         sviat         Pavms-S.pep         6*           NO:123         sviat         Pavms-S.pep         6           NO:124         sviat         Acams-S.pep         6     <	SEQ ID NO:103	sviat	Mims-C.pep	-	QVLSPQCQYGSIFWRNSYEHENMERLQCE
NO:105         sviat         Aams-2.pep         1           NO:106         sviat         Aams-4.pep         1           NO:107         sviat         Aams-6.pep         1           NO:108         sviat         PPd57-2ms.pep         1           NO:109         sviat         PPd57-3.pep         1           NO:110         sviat         PMms-B.pep         2           NO:111         sviat         PMms-E.pep         2           NO:112         sviat         PMms-D.pep         2           NO:114         sviat         PMms-C.pep         9           NO:115         sviat         PMms-C.pep         11           NO:116         sviat         PPMms-C.pep         9           NO:117         sviat         PPms-2.pep         6*           NO:120         sviat         Acams-4.pep         11           NO:121         sviat         Pavims-2.pep         6*           NO:122         sviat         Pavims-2.pep         6*           NO:123         sviat         Pavims-3.pep         6           NO:124         sviat         RTms-1.pep         6           NO:125         sviat         Acams-5.pep         6	<b>SEQ ID NO:104</b>	sviat	RTms-5.pep	τ-	QVLSPQCQYGSIFWRNSYEHENMERLQCE
NO:106         sviat         Aams-4.pep         1           NO:107         sviat         Aams-6.pep         1           NO:108         sviat         PPd57-2ms.pep         1           NO:109         sviat         PPMms-A.pep         2           NO:110         sviat         PMms-B.pep         2           NO:111         sviat         PMms-E.pep         2           NO:112         sviat         PMms-D.pep         2           NO:114         sviat         PMms-C.pep         9           NO:115         sviat         PMms-C.pep         13           NO:116         sviat         PPms-2.pep         14           NO:117         sviat         Pavms-4.pep         11           NO:118         sviat         Pavms-4.pep         6*           NO:120         sviat         Pavms-4.pep         6*           NO:121         sviat         Pavms-2.pep         6*           NO:122         sviat         Pavms-5.pep         6           NO:123         sviat         RTms-1.pep         6           NO:124         sviat         Acams-5.pep         6           NO:125         sviat         Acams-5.pep         6	SEQ 10 NO:105	sviat	Aams-2.pep	-	QVLSPQCQYGSIFWRNSYEHENMERLQCE
NO:107         sviat         Aams-6.pep         1           NO:108         sviat         PPd57-2ms.pep         1           NO:109         sviat         PPd57-3.pep         1           NO:110         sviat         PMms-B.pep         2           NO:111         sviat         PMms-E.pep         2           NO:112         sviat         PMms-D.pep         2           NO:114         sviat         PMms-D.pep         2           NO:115         sviat         PMms-C.pep         9           NO:116         sviat         PMms-C.pep         11           NO:117         sviat         PPavms-4.pep         11           NO:118         sviat         Pavms-4.pep         11           NO:120         sviat         Pavms-4.pep         6*           NO:121         sviat         Pavms-2.pep         6*           NO:122         sviat         Pavms-5.pep         6*           NO:123         sviat         Pavms-5.pep         6           NO:124         sviat         RTms-1.pep         6           NO:125         sviat         Acams-5.pep         6           NO:126         sviat         Acams-5.pep         6	SEQ ID NO:106	sviat	Aams-4.pep	Ψ-	<b>QVLSPQCQYGSIFWRNSYEHENMERLQCE</b>
NO:10B         svlat         PPd57-2ms.pep         1           NO:109         svlat         PPd57-3.pep         1           NO:110         sviat         PMms-A.pep         2           NO:111         sviat         PMms-E.pep         2           NO:112         sviat         PMms-D.pep         2           NO:114         sviat         PMms-D.pep         2           NO:115         sviat         PMms-C.pep         9           NO:116         sviat         PMms-C.pep         18*           NO:117         sviat         PMms-C.pep         14*           NO:118         sviat         PPMms-C.pep         14*           NO:119         sviat         Pavms-4.pep         11*           NO:120         sviat         Acams-2.pep         6*           NO:121         sviat         Pavms-2.pep         6*           NO:122         sviat         Pavms-5.pep         6           NO:123         sviat         Pavms-5.pep         6           NO:124         sviat         RTms-1.pep         6           NO:125         sviat         Acams-5.pep         6           NO:126         sviat         Acams-5.pep         6 <td><b>SEQ ID NO:107</b></td> <td>sviat</td> <td>Aams-6.pep</td> <td>τ-</td> <td>QVLSPQCQYGSIFWRNSYEHENMERLQCE</td>	<b>SEQ ID NO:107</b>	sviat	Aams-6.pep	τ-	QVLSPQCQYGSIFWRNSYEHENMERLQCE
NO:109         sviat         PPd57-3.pep         1           NO:110         sviat         PMms-A.pep         2           NO:111         sviat         PMms-B.pep         2           NO:112         sviat         PMms-E.pep         2           NO:113         sviat         PMms-D.pep         2           NO:114         sviat         PMms-D.pep         2           NO:115         sviat         PMms-C.pep         9           NO:116         sviat         PMms-C.pep         18*           NO:117         sviat         PPMms-C.pep         19*           NO:118         sviat         PPMms-C.pep         11*           NO:120         sviat         Pavms-4.pep         11*           NO:121         sviat         Pavms-2.pep         6*           NO:122         sviat         Pavms-2.pep         6*           NO:123         sviat         Pavms-5.pep         6           NO:124         sviat         RTms-1.pep         6           NO:125         sviat         Acams-5.pep         6           NO:126         sviat         Acams-5.pep         6           NO:127         sviat         Acams-5.pep         6	SEQ ID NO:108	svlat	PPd57-2ms.pep	-	QVLSPQCQYGSIFWRNSYEHENMERLQCE
NO:110         sviat         PMms-A.pep         2           NO:111         sviat         PMms-B.pep         2           NO:112         sviat         PMms-E.pep         2           NO:113         sviat         PPd57-4ms.pep         2           NO:114         sviat         PMms-D.pep         2           NO:115         sviat         PMms-C.pep         9           NO:116         sviat         PMms-C.pep         18*           NO:117         sviat         PPms-2.pep         14*           NO:118         sviat         Pavms-4.pep         11*           NO:120         sviat         Acams-4.pep         6*           NO:121         sviat         Pavms-2.pep         6*           NO:122         sviat         Pavms-2.pep         6*           NO:123         sviat         Pavms-5.pep         6           NO:124         sviat         RTms-1.pep         6           NO:125         sviat         Acams-5.pep         6           NO:126         sviat         Acams-5.pep         6           NO:127         sviat         Acams-5.pep         6           NO:128         sviat         Acams-5.pep         6	SEQ ID NO:109	sviat	PPd57-3.pep	₩.	QVLSPQCQYGSIFWRNSYEHENMERLQCE
NO:111         sviat         PMms-B.pep         2           NO:112         sviat         PMms-E.pep         2           NO:113         sviat         PPd57-4ms.pep         2           NO:114         sviat         PMms-D.pep         2           NO:115         sviat         PPd57-1ms.pep         12           NO:116         sviat         PMms-C.pep         9           NO:117         sviat         PPms-2.pep         14           NO:128         sviat         Pavms-4.pep         11           NO:129         sviat         Acams-4.pep         6*           NO:121         sviat         Pavms-2.pep         6*           NO:122         sviat         Pavms-2.pep         6*           NO:123         sviat         Pavms-5.pep         6           NO:124         sviat         RTms-1.pep         6           NO:125         sviat         Acams-5.pep         6           NO:126         sviat         Acams-5.pep         6           NO:127         sviat         Acams-5.pep         6           NO:128         sviat         Acams-5.pep         6           NO:129         sviat         Acams-5.pep         6 <td><b>SEQ ID NO:110</b></td> <td>sviat</td> <td>PMms-A.pep</td> <td>73</td> <td>QVLSPQSQYGSIYWRNSYENENMERLQCE</td>	<b>SEQ ID NO:110</b>	sviat	PMms-A.pep	73	QVLSPQSQYGSIYWRNSYENENMERLQCE
NO:112         sviat         PMms-Epep         2           NO:113         sviat         PPd57-4ms.pep         2           NO:114         sviat         PMms-D.pep         2           NO:115         svlat         PPd57-1ms.pep         12           NO:116         svlat         PMms-C.pep         9           NO:117         svlat         PPms-2.pep         18*           NO:118         sviat         PPms-4.pep         11           NO:120         sviat         Acams-4.pep         11           NO:121         sviat         Acams-2.pep         6*           NO:122         sviat         Pavims-2.pep         6*           NO:123         sviat         Pavims-5.pep         6           NO:124         sviat         RTms-1.pep         6           NO:125         sviat         Acams-5.pep         6           NO:126         sviat         Acams-5.pep         6           NO:127         sviat         Acams-5.pep         6           NO:128         sviat         Acams-5.pep         6           NO:129         svixt         RTms-6.pep         6           NO:129         svixt         RTms-6.pep         6 <td>SEQ ID NO:111</td> <td>sviat</td> <td>PMms-B.pep</td> <td>7</td> <td>QVLSPQSQYGSIYWRNSYENENMERLQCE</td>	SEQ ID NO:111	sviat	PMms-B.pep	7	QVLSPQSQYGSIYWRNSYENENMERLQCE
NO:113         sviat         PPd57-4ms.pep         2           NO:114         sviat         PMms-D.pep         2           NO:115         sviat         PPd57-1ms.pep         12           NO:116         sviat         PMms-C.pep         9           NO:117         sviat         PPms-2.pep         19*           NO:128         sviat         PPms-4.pep         11           NO:129         sviat         Acams-4.pep         11           NO:121         sviat         Pavms-2.pep         6*           NO:122         sviat         Pavms-2.pep         6*           NO:123         sviat         Pavms-5.pep         6           NO:124         sviat         Pavms-5.pep         6           NO:125         sviat         RTms-1.pep         6           NO:126         sviat         Acams-5.pep         6           NO:127         sviat         Acams-5.pep         6           NO:127         sviat         Acams-5.pep         6           NO:128         sviat         Acams-5.pep         6           NO:129         svixt         RTms-6.pep         6           NO:129         svixt         RTms-6.pep         6 <td>SEQ ID NO:112</td> <td>sviat</td> <td>PMms-E.pep</td> <td>7</td> <td>QVLSPQSQYGSIYWRNSYENENMERLQCE</td>	SEQ ID NO:112	sviat	PMms-E.pep	7	QVLSPQSQYGSIYWRNSYENENMERLQCE
NO:114         sviat         PMms-D.pep         2           NO:115         svlat         PPd57-1ms.pep         12           NO:116         svlat         PMms-C.pep         9           NO:117         svlat         PMms-C.pep         18*           NO:118         sviat         PPms-2.pep         11*           NO:120         sviat         Acams-4.pep         11           NO:121         sviat         Pavms-2.pep         6*           NO:122         sviat         Pavms-2.pep         6*           NO:123         sviat         Pavms-2.pep         6*           NO:124         sviat         Pavms-5.pep         6           NO:125         sviat         RTms-1.pep         6           NO:126         sviat         Acams-5.pep         6           NO:127         sviat         Acams-5.pep         6           NO:127         sviat         Acams-5.pep         6           NO:128         sviat         RTms-6.pep         6           NO:129         svixt         RTms-6.pep         6           NO:129         svixt         RTms-6.pep         6	SEQ ID NO:113	sviat	PPd57-4ms.pep	7	QVLSPQSQYGSIYWRNSYENENMERLQCE
NO:115         svlat         PPd57-1ms.pep         12           NO:116         svlat         PMms-C.pep         9           NO:117         svlat         LGAms-6.pep         18*           NO:118         sviat         PPms-2.pep         11           NO:120         sviat         Acams-4.pep         11           NO:121         sviat         Acams-2.pep         6*           NO:122         sviat         Pavms-2.pep         6*           NO:123         sviat         Pavms-5.pep         6           NO:124         sviat         LGAms-5.pep         6           NO:125         sviat         Acams-3.pep         6           NO:126         sviat         Acams-3.pep         6           NO:127         sviat         Acams-5.pep         6           NO:128         sviat         Acams-5.pep         6           NO:129         sviat         RTms-6.pep         6           NO:129         sviat         RTms-6.pep         6           NO:129         sviat         RTms-6.pep         6	SEQ ID NO:114	sviat	PMms-D.pep	2	QVLSPQSQYGSIYWRNSYENENMERLQCE
NO:116         svlat         PMms-C.pep         9           NO:117         svlat         LGAms-6.pep         18*           NO:118         svlat         PPms-2.pep         11           NO:120         svlat         Pavms-4.pep         11           NO:121         svlat         Acams-4.pep         6*           NO:122         svlat         Pavms-2.pep         6*           NO:123         svlat         Pav5ms.pep         6           NO:124         svlat         LGAms-5.pep         6           NO:125         svlat         RTms-1.pep         6           NO:126         svlat         Acams-37.pep         6           NO:127         svlat         Acams-5.pep         6           NO:128         svlat         Acams-5.pep         6           NO:129         svvat         RTms-6.pep         6           NO:129         svvat         RTms-6.pep         6           NO:129         svsat         RTms-2.pep         6	SEQ (D NO:115	sviat	PPd57-1ms.pep	12	QVLSPQTQYGSYWRNSYENENMERLQCE
VO:117         svlat         LGAms-6.pep         18*           VO:118         svlat         PPms-2.pep         19*           VO:120         svlat         Pavms-4.pep         11           VO:121         svlat         Acams-4.pep         6*           VO:122         svlat         Pavms-2.pep         6*           NO:122         svlat         Pav5ms.pep         6           NO:123         svlat         RTms-1.pep         6           NO:125         svlat         RTms-1.pep         6           NO:126         svlat         Acams-3.pep         6           NO:127         svlat         Acams-5.pep         6           NO:128         svlat         Acams-5.pep         6           NO:129         svvat         RTms-6.pep         6           NO:129         svvat         RTms-6.pep         6           NO:129         svsat         RTms-2.pep         6	SEQ ID NO:116	sviat	PMms-C.pep	o	<b>QVLSPQTQYGSIYWRNSYENGNMERLQCE</b>
sviat PPms-2.pep 19* sviat Pavms-4.pep 11 sviat Acams-4.pep 3 sviat Acams-2.pep 6* sviat Pavms-2.pep 6* sviat Pavms-5.pep 6 sviat RTms-1.pep 6 sviat RTms-1.pep 6 sviat Acams-5.pep 6 sviat RTms-6.pep 6	SEQ ID NO:117	sviat	LGAms-6.pep	18*	<b>QVLSPQYQYGSIFWRNSYENENMERLQCE</b>
sviat Pavms-4.pep 11 sviat Acams-4.pep 3 sviat Acams-2.pep 6* sviat Pavms-2.pep 6* sviat Pavms-5.pep 6 sviat RTms-1.pep 6 sviat RTms-1.pep 6 sviat Acams-3.pep 6 sviat Acams-5.pep 6 sviat RTms-6.pep 6	SEQ 10 NO:118	sviat	PPms-2.pep	19*	<b>QVLSPQYQYGSIFWRNSYENENMERLRCE</b>
sviat Acams-4.pep 3 sviat Acams-2.pep 6* sviat Pavms-2.pep 6* sviat Pav5ms.pep 6 sviat LGAms-5.pep 6 sviat RTms-1.pep 6 sviat RTms-1.pep 6 sviat Acams-3.pep 6 sviat Acams-5.pep 6 sviat Acams-5.pep 6 sviat Acams-5.pep 6 sviat Acams-5.pep 6 sviat RTms-6.pep 6	SEQ ID NO:119	sviat	Pavms-4.pep	7	QVLSPQYQYGSIYWGNSYENENMERLQCE
sviat Acams-2.pep 6* sviat Pavms-2.pep 6* sviat Pav5ms.pep 6* sviat LGAms-5.pep 6 sviat RTms-1.pep 6 sviat RTms-1.pep 6 sviat Acams-3.pep 6 sviat Acams-5.pep 6 sviat Acams-5.pep 6 sviat Acams-5.pep 6 sviat RTms-6.pep 6 sviat RTms-6.pep 6	SEQ ID NO:120	sviat	Acams-4.pep	თ	<b>QVLSPQYQYGSIYWRNSHENENMERLQCE</b>
sviat Pavms-2.pep 6* sviat Pav5ms.pep 6 sviat LGAms-5.pep 6 sviat RTms-1.pep 6 sviat RTms-1.pep 6 sviat Acams-3.pep 6 sviat Acams-5.pep 6 svivt RTms-6.pep 6 svsvit RTms-6.pep 6	SEQ (D NO:121	sviat	Acams-2.pep	<b>0</b> *	QVLSPQYQYGSIYWRNSYENENMERLQCE
svlat Pav5ms.pep 6 svlat LGAms-5.pep 6 svlat RTms-1.pep 6 svlat Pavms-3.pep 6 svlat Acams-3.pep 6 svlat Acams-5.pep 6 svlat Acams-5.pep 6 svlat RTms-6.pep 6 svsat RTms-6.pep 6	SEQ ID NO:122	sviat	Pavms-2.pep	ů,	QVLSPQYQYGSIYWRNSYENENMERLQCE
sviat LGAms-5.pep 6 sviat RTms-1.pep 6 sviat Pavms-3.pep 6 sviat Acams-37.pep 6 svlat Acams-5.pep 6 svivt RTms-6.pep 6 svsat RTms-2.pep 6	SEQ ID NO:123	sviat	Pav5ms.pep	ø	QVLSPQYQYGSIYWRNSYENENMERLQCE
sviat RTms-1.pep 6 sviat Pavms-3.pep 6 sviat Acams-37.pep 6 svlat Acams-5.pep svivt RTms-6.pep 6 svsat RTms-2.pep 6	SEQ ID NO:124	sviat	LGAms-5.pep	9	QVLSPQYQYGSIYWRNSYENENMERLQCE
sviat Pavms-3.pep 6 sviat Acams-3?.pep 6 svlat Acams-5.pep svivt RTms-6.pep 6 svsat RTms-2.pep 6	SEQ ID NO:125	sviat	RTms-1.pep	ဖ	QVLSPQYQYGSIYWRNSYENENMERLQCE
sviat Acams-37.pep 6 svlat Acams-5.pep svivt RTms-6.pep 6 svsat RTms-2.pep 6	SEQ ID NO:126	sviat	Pavms-3.pep	9	QVLSPQYQYGSIYWRNSYENENMERLQCE
svlat Acams-5.pep 6 svivt RTms-6.pep 6 svsat RTms-2.pep 6	<b>SEQ ID NO:127</b>	sviat	Acams-37.pep	9	QVLSPQYQYGSIYWRNSYENENMERLQCE
svivt RTms-6.pep 6 svsat RTms-2.pep 6	SEQ ID NO:128	svlat	Acams-5.pep		
svsat RTms-2.pep 6	SEQ ID NO:129	svivt	RTms-6.pep	9	<b>QVLSPQYQYGSIYWRNSYENENMERLQCE</b>
	SEQ ID NO:130	svsat	RTms-2.pep	9	QVLSPQYQYGSIYWRNSYENENMERLQCE

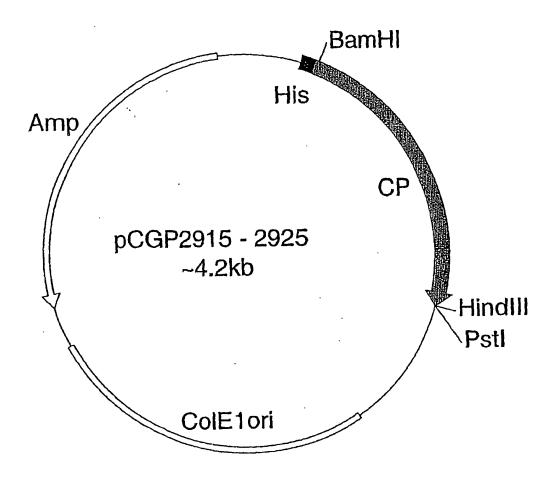
# Figure 9 continued

, A

variable amino acids across sequence

SEQ ID NO:92 SEQ ID NO:93

## Figure 9 continued



Replicon: pQE30 BamHI/PstI ~3.5kb vector

Insert: ~0.7kb <u>BamHI/PstI</u> PCR products generated using visproF1 and visproR1 primers and cDNA prepared from RNA isolated from various marine organisms

Figure 10

Rtms5,  $\varepsilon_{592}$  = 111,000 M<sup>-1</sup>cm<sup>-1</sup>

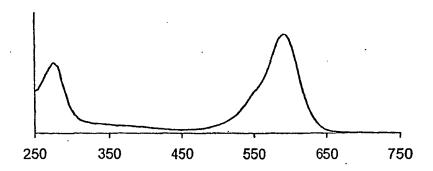


Figure 11(a)

LGAsv-C  $\epsilon_{591} = 53,000 \text{ M}^{-1} \text{cm}^{-1}$ 

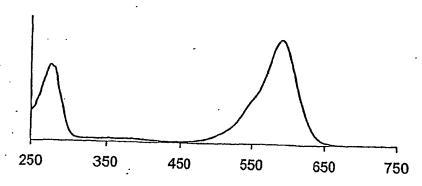


Figure 11(b)

Ce61-7sv;  $\varepsilon_{s91.5} = 104,000 \text{ M}^{-1} \text{cm}^{-1}$ 

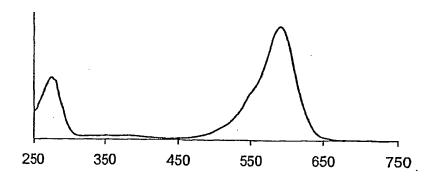


Figure 11(c)

PPd57-2ms,  $\varepsilon_{593} = 67,000 \text{ M}^{-1} \text{cm}^{-1}$ 

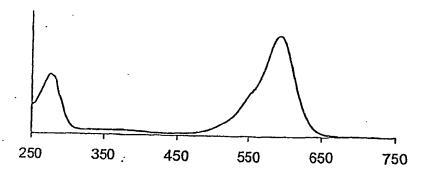


Figure 11(d)

MimsC,  $\varepsilon_{589} = 48,000 \text{ M}^{-1} \text{cm}^{-1}$ 

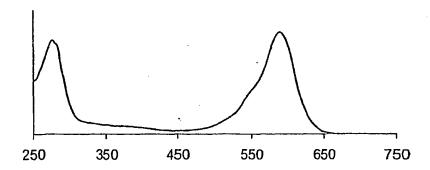
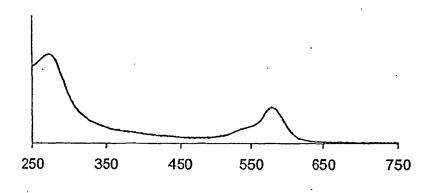


Figure 11(e)

PMms-B,  $\varepsilon_{579.5} = 39,000 \text{ M}^{-1} \text{cm}^{-1}$ 



rtsv2  $\epsilon_{579.5} = 75,000 \text{ M}^{-1} \text{cm}^{-1}$ 

### Figure 12A(a)

LGAsv-D,  $\varepsilon_{579} = 72,400 \text{ M}^{-1} \text{cm}^{-1}$ 

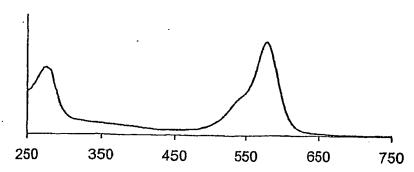


Figure 12A(b)

rtsv2  $\epsilon_{579.5}$  = 75,000 M<sup>-1</sup>cm<sup>-1</sup>

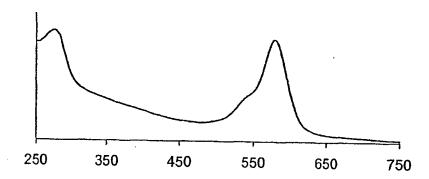


Figure 12A(c)

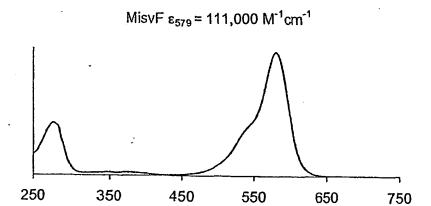


Figure 12B(d)

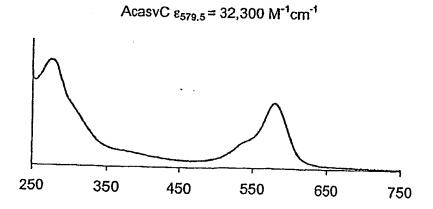


Figure 12B(e)

**30/94** 

LGAms5  $\epsilon_{583.5} = 71,000 \text{ M}^{-1} \text{cm}^{-1}$ 

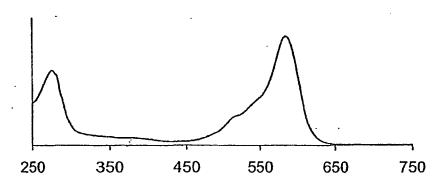


Figure 13(a)

Rtms1  $\varepsilon_{584} = 44,000 \text{ M}^{-1} \text{cm}^{-1}$ 

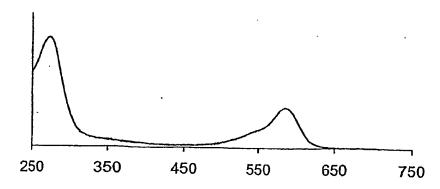
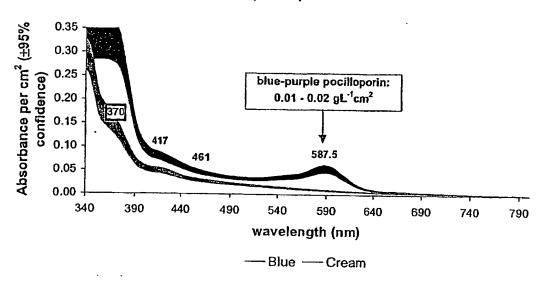


Figure 13(b)





### Figure 14(a)

### Chromatogram of gel filtrated protein elution

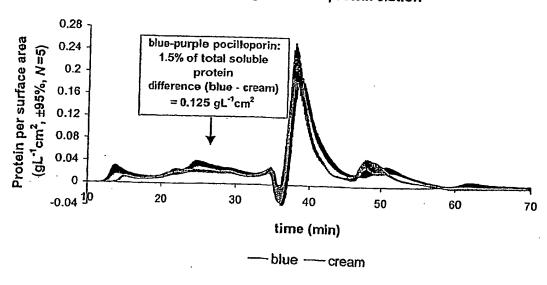
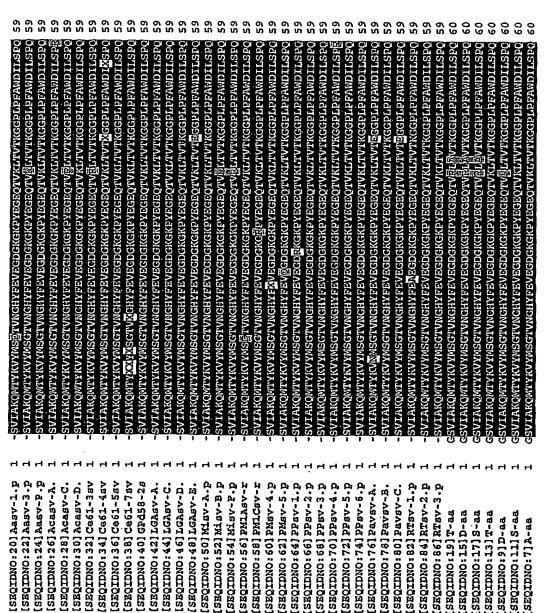


Figure 14(b)

T1-aa	1	GSVIAKQMTYKVYMSGTVNGHYFEVEGDGKGKPYEGEQTVRLAVTKGGPL	50
D1-aa	1		50
S1-aa	1	GSVIAKQMTYKVYMSGTVNGHYFEVEGDGKGKPYEGEQTVRLAVTKGGPL	50
T3-aa		GSVIAKQMTYKVYMSGTVNGHYFEVEGDGKGKPYEGEQTVKLTVTKGGPL	50
D10-aa	1	GSVIAKQMTYKVYMSGTVNGHYFEVEGDGKGKPYEGEQTVRLTVTKGGPL	50
S3-aa	1	GSVIAKQMTYKVYMSGTVNGHYFEVEGDGKGKPYEGEQTVKLTVTKGGPL	50
A8-aa		GSVIAKQMTYKVYMSGTVNGHYFEVEGDGKGKPYEGEQTVKLTVTKGGPL	50
		*************	
Tl-aa	51	PFAWDILSPQCQYGSIPFTKYPEDIPDYVKRSFPEGFTWERIMNFEDGAV	100
D1-aa	51	PFAWDILSPQCQYGSIPFTKYPEDIPDYVKQSFPEGFTWERIMNFEDGAV	100
S1-aa	51		100
T3-aa	51	PFAWDILSPQSQYGSIPFTKYPEDIPDYVKQSFPEGYTWERIMNFEDGAV	100
D10-aa	51	PFAWDILSPQSQYGSIPFTKYPEDIPDYVKQSFPEGYTWERIMNFEDGAV	100
S3-aa	51	PFAWDILSPQSQYGSIPFTKYPEDIPDYVKQSFPEGYTWERIMNFEDGAV	100
A8-aa	51	PFAWDILSPQSQYGSIPFTKYPEDIPDYVKQSFPEGYTWERIMNFEDGAV	100
		****************	
T1-aa	101	CTVSNDSSIQGNCFIYHVKFSGLNFPPNGPVMQKKTQGWEPHSERLFARD	150
D1-aa		CPVSNDSSIQGNCFIYHVKFSGLNFPPNGPVMQKKTQGWEPHSERLFARD	150
S1-aa		CTVSNDSSIQGNCFIYHVKFSGLNFPPNGPVMQKKTQGWEPNTERLFARD	150
T3-aa	101	CTVSNDSSIQGNCFIYHVKFSGLNFPPNGPVMQKKTQGWEPNTERLFARD	150
D10-aa		CTVSNDSSIQGNCFIYHVKFSGLNFPPNGPVMQKKTQGWEPNTERLLARD	150
S3-aa		CTVSNDSSIQGNCFIYHVKFSGLNFPSNGPVMQKKTQGWEPNTERLFARD	150
A8-aa	101	CTVSNDSSIQGNCFIYHVKFSGLNFPPNGPVMQKKTQGWEPNTERLFARD	150
		* ***************** **********	
T1-aa	151	GMLIGNNFMALKLEGGGHYLCBFKTTYKAKKPVKMPGYHYVDRKLDVINH	200
D1-aa	151	GMLIGNTFMALKLEGGGHYLCEFKTTYKAKKPVKMPGYHYVDRKLDVINH	200
S1-aa	151	GMLIGNNFMALKLEGGGHYLCEFKSTYKAKKPVKMPGYHYVDRKLDVTNH	200
T3-aa		GMLIGNNFMALKLEGGGHYLCEFKSTYKAKKPVKMPGYHYVDRKLDVTNH	200
D10-aa		GMLIGNNFMALKLEGGGHYLCEFKSTYKARKPVKMPGYHYVDRKLDVTNH	
S3-aa		GMLIGNNFMALKLEGGGHYLCEFKSTYKAKKPVKMPGYHYVDRKLDVTNH	
A8-aa	151	A TO TAKE A TOTAL AND A TIME A TOTAL AND A TIME	200
		******.***************************	
T1-aa		NKDYTSVEQCEISIARKPVVALQ 223	
D1-aa	201		
S1-aa		NKDYTSVEQCEISIARKPLVALQ 223	
T3-aa		NKDYTSVEQCEISIARKPVVALQ 223	
D10-aa		NKDYTSVEQREISIARKPVVALQ 223	
S3-aa		NKDYTSVEQCEISIARKPLVALQ 223	
A8-aa	20I	NKDYTSVEQCEISIARKPVVALQ 223	
		******* ***** ***	



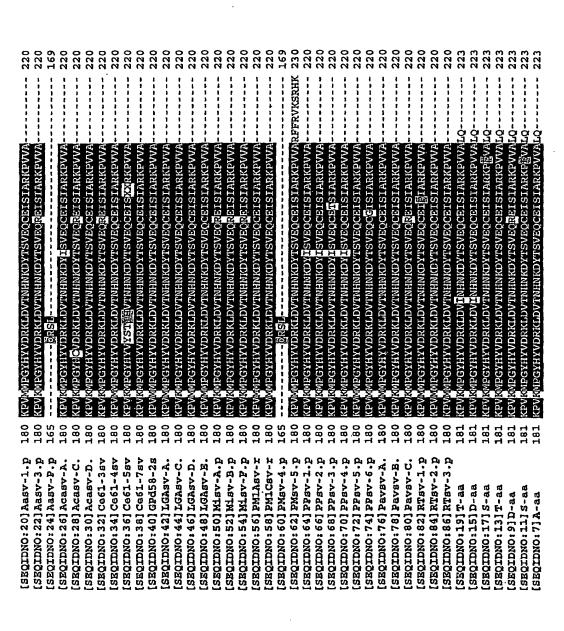
### Figure 16

# Figure 16 continued

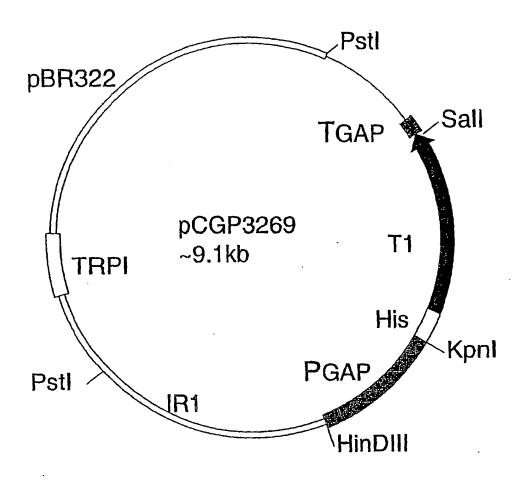
[SEQIDNO: 20] Assv-1.p	90	Name.
[SECIDNO:24] Assv-P.p	9 0	SOXEST PETIKK PEDITEDIK VIKASE PEGERMEKTIMINE BUGAN CITVSINDSSIQGINCE LIGAN FILE SOXEST PETIKK PEDITEDIK VIKASE PEGERMERITÄN PEDGAN CITVSINDSSIOGINCET VINVIKE 11.9
[SEQIDNO: 26] Acasv-A.	9	HETWERT WINSEDGAVCTVSNDSSTOGNOFT VEIVER
[SEQIDNO: 28] Acasv-C.	80	IPDYVKOSFPEGATWERIMNFEDGAVCTVSNDSSIOGNCFIYE
[SEQIDNO: 30] Acasv-D.	9	FINERIMNFEDGAVCTVSNDSSIQGNCFTYFIVE
[SEQIDNO:32]Ce61-3sv	9	#TWERIMNFEDGAVCTVSNDSSIQGNCFIY#JVKF
[SECIDNO:34] Ce61-4sv	g (	MTWERIMNFEDGAVCTVSNDSSIQGNCFIYEVE
SECTIONO: 381 CAST - 28A	ο ς ν φ	OVESTAPLINY PEDAPONANA SEPECTATION OF THE TRANSPORT OF TH
[SECIDIO: 40] GPGS8-23	9	MATERIAN PEDICAY CITY SAIDSSTOCK TANAKE
[SEQIDNO: 42] LGASV-A.	9	MTWERTWINFEDGAVCTVSNDSSTOGNCFIYNVKF
[SECIDNO: 44] LGABV-C.	9	TWERTWINFEDGAVCTVSNDSSTOGNCFTYRVKF
(SECIDNO: 46] LGASV-D.	9	THERIMNFEDGAVCTVSNDSSIQGNCFIYNVKF
[SECIDNO: 48] LGABV-E.	9	s <mark>k</mark> twertwnfedgavctvsndsstogncftykvkf
[SECIDNO: 50] Misv-A.p	9	BETWERIMNFEDGAVCTVSNDSSIQGNCFIYEVKF
[SEQIDNO: 52] Misv-B.p	9	3 THERIMNFEDGAVCTVSNDSSIQGNCFIX OVE
[SECIDNO: 54] MIBV-F. D	9	WIMERIMNFEDGAVCTVSNDSSIQGNCFIYENKF
[SECTONO: 56] PMIASV-T	8	SHIWERIMMFEDGAVCTVSNDSSIQGNCFIYNVKF
[SECIDNO:58] PMICBV-r	0	SHTRERIMEFEDGAVCTVSNDSSHOGNCFIYNVKF
[SECIDNO: 60] PMav-4.p	90	TWERT WINFEDGAVCTVSNDSSTOGNCFTYNNKF
[SECIDNO: 62] PMSV-5.p	9	MTWERIMNFKDGAVCTVSNDSSI QGNCFI YMVKF
[SECIDNO: 64] PPSV-1.p	9	IPFTKYPEDMPDYVKQSFPEGETWERIMNFEDGAVCTVSNDSSIQGNCFTYMUTT
[SECIDNO: 66] PPBV-2.p	9	ETWERIMNFEDGAVCTVSNDSSIQGNCFTYEVEF
[SEQIDIO: 68] PP8v-3.p	00 (	IPFTKYPEDEPDYVKOSFPECETWERIWNFEDGAVCTVSNDSSIQGNCFTYEVKF
SECLDING: 70] PPSV-4.D	9 (	LPFTKZPELINPDXVKOSPPEGETWERIMNFEDGAVCTVSNDSSTOGNOFTYEVKF
[SEQIDNO: 72] PPav-5.p	8	TWERIMNFEDGAVCTVSNDSSIDGNCFIYNVKF 11
CAROLLINO: /4] FEBV-6.D	2 (	ELWERT WINE EDGAVCTVSNDSSLCGGCFT XINV KA
CAPOTTONO: 781 Daviers B	ט פ	SALESSAR TATES TRUE PLANCE FEGRE FRESHINKE ELGANCIVENIUS SALES TANDENIUS SALES TATES TO SALES TATES TO SALES TA
SECIDIO: 80] Payer-C.	9	STABLE IMMEDIATION OF THE STATE
[SECIDNO: 82] RISV-1.p	09	WITHER HANDED GAVOTVSNDSSTOGNOFT XNVKF
[SECIDNO: 84] RIBV-2.p	9	SQYGSIPFTKYPEDIPDYVKQSFPEGHTWERIMNFEDGAVCTVSNDSSIQGNCFIYMVIF 119
[SEQIDNO:86] RIBV-3.p	9	SOYGSIPFTKYPEDIPDYVKOGFPEGÄTMERIMNFEDGAVCTVSNDSSIQGNCFIYÄÄVKF 119
[SECIDNO:19] T-aa	61	Coxestpetikypedtedyvaräsppegätmerimnfedgavctvsndssiogncfivävkf 120
(SEQIDNO:15]D-aa	61	FIMERIMNEEDGAVCPVSNDSSIOGNCFIYEVKF
[SECIDNO: 17] S-aa	; 6	ETWERIMNFEDGAVCTVSNDSSIQGNCFIYEVE
[SECIDNO: 13] T-aa	61	MINERIMNFEDGAVCTVSNDSSIQGNCFIYAVKF
[SEQIDNO: 9]D-82	19 ;	THE BRUNNIE BUCKNOWN SELECTION OF THE WAY
(SECTING: 11) S-BB	7 5	SQXGSIPFTKXYBEDIEDIYVKQSFPBGGTWERIWNFBBGAVCTVSWBSSIQGWCFIYGVKF 120
	<b>†</b>	THE THE PROPERTY OF THE PROPER

### 179 179 179 179 179 179 179 179 179 180 180 180 179 179 179 179 179 179 179 179 165 179 179 179 180 TYKAE TYKAE TYKAE TYKAK TYKAK TYKAK TYKAK TYKAK TYKAK TYKAK VYKAK IVKA IVKAK IVKAK TYKAK TYKAK TYKAK YKKAK TYKAK TYKAK TYKAK YKAK YKANK PYKAK PYKAK MERLIPARDGMLIGNNEWALKLEGGGHYLCEFKETY MERLIPARDGMLIGNNEWALKLEGGGHYLCEFKETY MERLIPARDGMLIGNNEWALKLEGGGHYLCEFKETY MERLIPARDGMLIGNNEWALKLEGGGHYLCEFKETY MERLIPARDGMLIGNNEWALKLEGGGHYLCEFKETY MERLIPARDGMLIGNNEWALKLEGGGHYLCEFKETY MERLIPARGGMLIGNNEWALKLEGGGHYLCEFKETY MERLIPARGGMLIGNNEWALKLEGGGHYLCEFKETY MERLIPARGGMLIGNNEWALKLEGGGHYLCEFKETY MERLIPARGGMLIGNNEWALKLEGGGHYLCEFKETY MERLIPARGGMLIGNNEWALKLEGGGHYLCGFKETY MERLIPARGGMLIGNNEWALKLEGGGHYLCGFKETY MERLIPARGGMLIGNNEWALKLEGGGHYLCGFKETY ERLFAREGMLIGNNPWALKLEGGGHYLCGFREN BERLFAREGMLIGNNPWAFKLEGGGHYLCEPKEN BERLFARDGMLIGNNPWALKLEGGGHYLCEPKEN ERLFARDGMLIGNNPWALKLEGGGHYLCEPKEN ERRJFARDGMLIGNNPWALKTGGGGHYLCEFKEN ERRJFAREGMLIGNNPWALKLEGGGHYLCEFKEN SERLFARDGWLIGNNFWALKLEGGCHYLCEFREN BERLFARDGWLIGNNFWALKLEGGCHYLCEFREN BERLFARDGWLIGNNFWALKLEGGCHYLCEFREN BERLFARDGWLIGNNFWALKLEGGCHYLCEFREN BERLFARDGWLIGNNFWALKLEGGCHYLCEFREN BERLFARDGWLIGNNFWALKLEGGCHYLCEFREN BERLFARDGWLIGNNFWALKLEGGGHYLCEFKE BERLFARDGWLIGNNFWALKLEGGGHYLCEFKE SERLFARGGWLIGNNFWALKLEGGGHYLCEFKE WERLFARDGWLIGNNFWALKLEGGGHYLCEFKE BRLFARDGWLIGNNFWALKLEGGGHYLCEFKE BRLFARDGWLIGNNFWALKLEGGGHYLCEFKE erleardgmlignnfwalkleggghylceff erlearggmlignnfwalklèggghylceff erleardgmlignnfwalklèggghylceff erleardgmlignnfwalkleggghylceff erlfardgwlignnfmalkieggghylcef! srlfardgmetgnnfwalkebegghylger erleardgwlignnfwalkleggghylcef SGINFPPNGPVWQKKTQGWEPN SGINFPPNGPVWQKKTQGWEPN SGINFPPNGPVWQKKTQGWEPN SGINFPPNGPVWQKKTQGMEPN SGINFPPNGPVWQKKTQGMEPN SGLNFPPNGPVMQKKTQGWEPIK SGLNFPPNGPVMQKKTQGWEPIK SGLNFPPNGPVMQKKTQGWEPIK SGLNFPPNGPVMQKKTQGWQPIK SGLNFPPNGPVMQKKTQGMQPIK SGLNFPPNGRVMQKKTQGNEP **BELINEPPINGPVMOKKTOGMEP GENNPPINGPVMQKKCYQGWEP** SCLNFPPNGPVMQKKTQCWEP sglnpppngpvmokktoggnbf 3GLNFPPNGPVMQKKTQGWEP SGLNFPPNGPVWQKKTQGWEI SGLNFPPNGPVWQKKTQGWEI SGLNFPPNGPVWQKKTQGWEI 3GLNFPPNGPVMQKKTQGWER SGLNFPPNGPVMQKKTQGWEB SGLNFPPNGPVMQKKTQGMEI SGLWFPPNGPVMQKKTQGMEI SGLNFPPNGPVMQKKTQGMEI **SCLNFPPNGPVMOKKTQGWE**I nfpengngpvmokktogne) SGINFPPNGPVMQKKTQGM) SGINFPPNGPVMQKKTPGGM) SGINFPPNGPVMQKKTQGM) Scinifppngpvmokktogm **GLNFPPNGPVMQKKTQGW** icenpppngpvmokktogs 120 120 120 120 120 120 120 120 1120 1120 1120 1120 1120 1120 1120 SECIDNO:84] RISV-2.D (SEQIDNO: 20] Aasv-1.p SECIDNO: 34] Ce61-48v SEQIDNO:36] Ce61-5av [SEQIDNO:54] Misv-F.p. (SEQIDNO:56] PMIASV-r SEQIDNO: 60] PMSV-4.p [SEQIDNO: 62] PMSV-5.p (SEQIDNO: 64] PPBV-1.p SECIDNO: 661 PPsv-2.p (SEQIDNO: 68) PPsv-3.p SECIDNO:72] PPsv-5.p SECIDNO:74] PPsv-6.p SEQIDNO: 24] Aasv-P.p SEQIDNO:32] Ce61-38v SEQIDNO:38] Ce61-7sv [SEQIDNO:40] GPd58-28 SEQIDNO:50] Misv-A.p SEQIDNO: 52] Misv-B.p SEQIDNO: 58] PMLCsv-r SEQIDNO:70] PPsv-4.p SEQIDNO:82] RISV-1.p [SEQIDNO: 26] Acasv-A. SEQIDNO: 42] LGASV-A. SECIDNO: 44] LGASV-C. SEQIDNO: 46] LGASV-D. SEQIDNO:76] Pavsv-A. SEQIDNO: 28] Acasv-C. [SECIDNO: 30] Acaby-D. SEQIDNO: 48] LGASV-E. SEQIDNO:78] Pavev-B. SECIDNO: 80] Pavsv-C. SECTONO: 19] T-aa SEQIDNO:15]D-aa SECIDNO:17] S-aa SEQIDNO: 13] T-aa SECIDNO: 11] 8-aa SECIDNO: 9] D-aa SEQIDNO:71A-aa

# Figure 16 continued



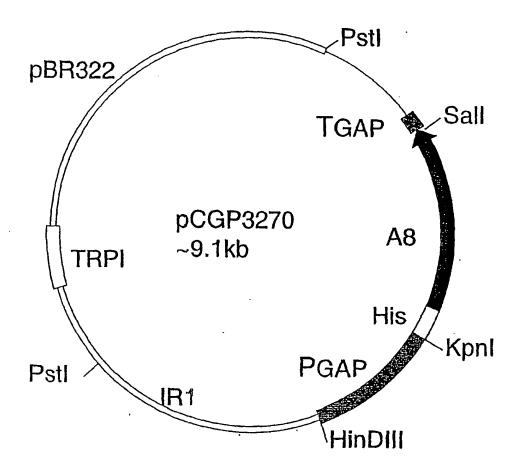
# Figure 16 continued



Replicon: pYE22m Kpnl/Sall

Insert: ~0.7kb Kpnl/Sall PCR products generated using Kpn.6His.F and T1/A8.Sal.R primers and pCGP2921 as template

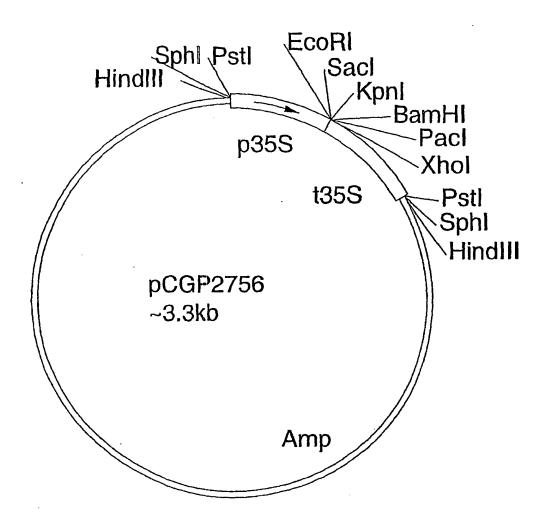
Figure 17



Replicon: pYE22m Kpnl/Sall

Insert: ~0.7kb Kpnl/Sall PCR products generated using Kpn.6His.F and T1/A8.Sal.R primers and pCGP2918 as template

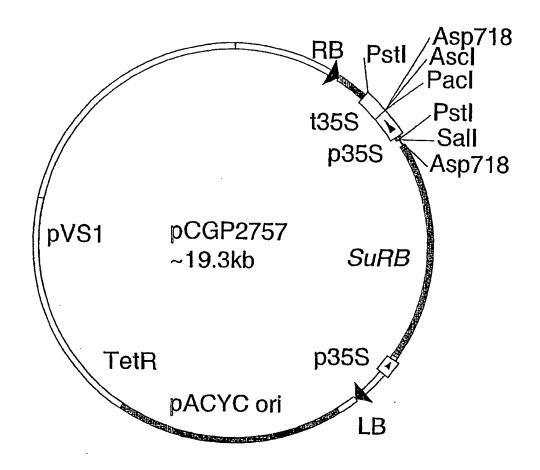
Figure 18



Replicon: pRTppoptc EcoRl/Xbal 3.3kb vector

Insert: ~40bp <u>EcoRI/Xbal</u> fragment containing multiple cloning site from pNEB193

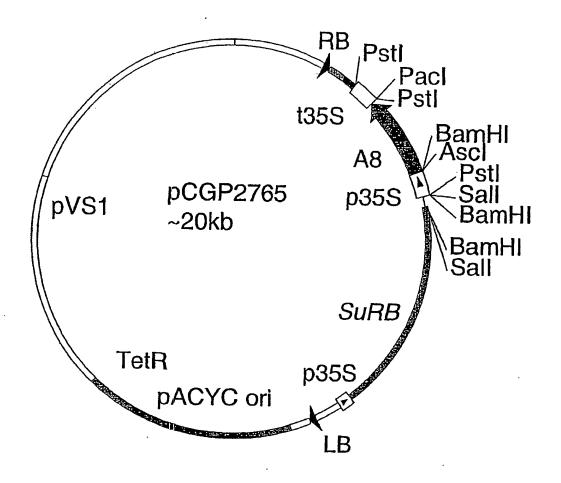
Figure 19



Replicon: pWTT2132 Pstl ~18.6kb vector

Insert: ~0.7kb Pstl fragment from pCGP2756

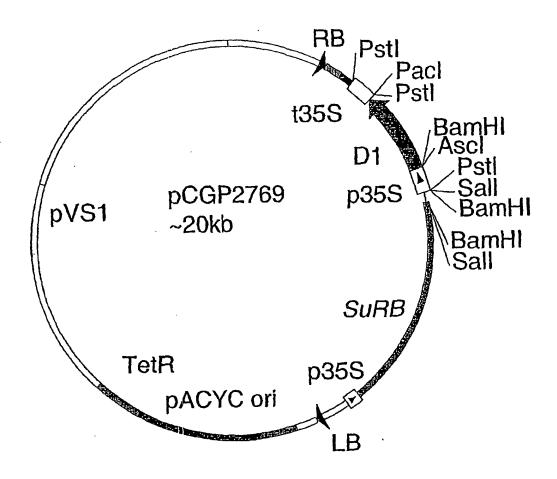
Figure 20



Replicon: pCGP2757 Ascl/Pacl ~19.3kb vector

Insert: ~0.7kb <u>Ascl/Pacl</u> A8 PCR product using visproF1 and visproR1 primers and pCGP2918 as template

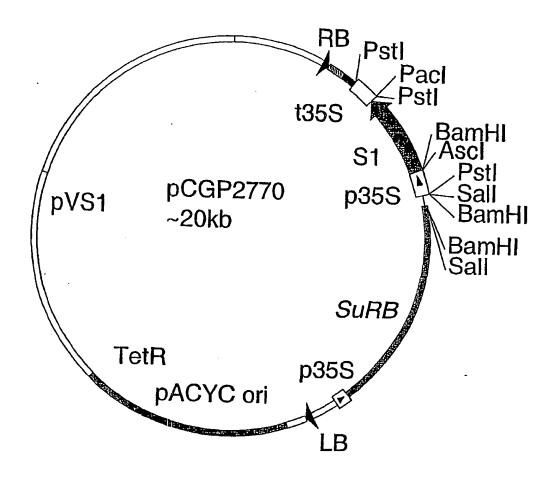
Figure 21



Replicon: pCGP2757 Ascl/Pacl ~19,3kb vector

Insert: ~0.7kb <u>Ascl/Pacl</u> D1 PCR product using visproF1 and visproR1 primers and pCGP2919 as template

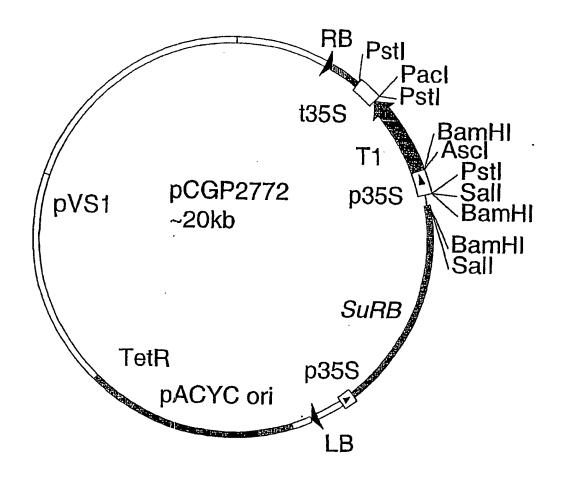
Figure 22



Replicon: pCGP2757 Ascl/Pacl ~19.3kb vector

Insert: ~0.7kb <u>Ascl/Pacl</u> S1 PCR product using visproF1 and visproR1 primers and pCGP2923 as template

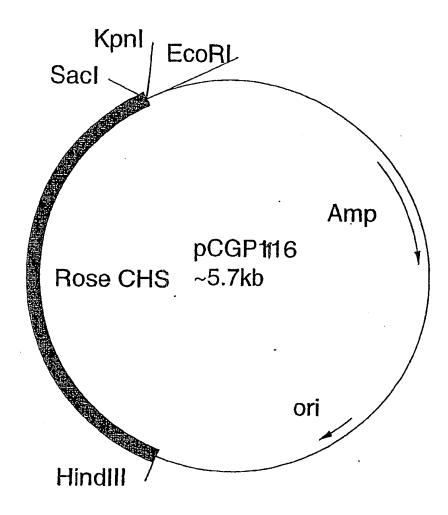
Figure 23



Replicon: pCGP2757 Ascl/Pacl ~19.3kb vector

Insert: ~0.7kb <u>Ascl/Pacl</u> T1 PCR product using visproF1 and visproR1 primers and pCGP2921 as template

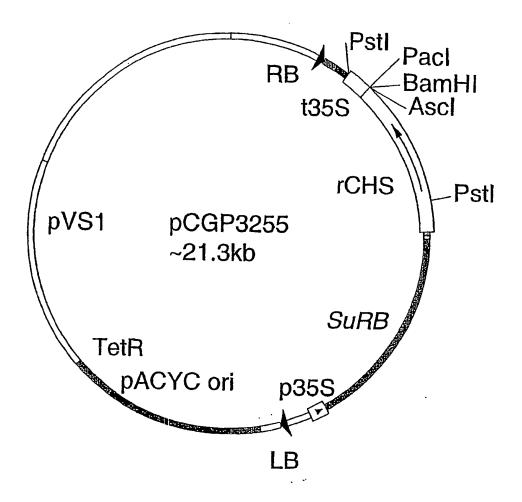
Figure 24



Replicon: pUC19 HindlII/Smal 2.7kb vector

Insert: <u>HindIII/EcoRV</u> fragment from pCGP1114 containing the Rose CHS promoter fragment

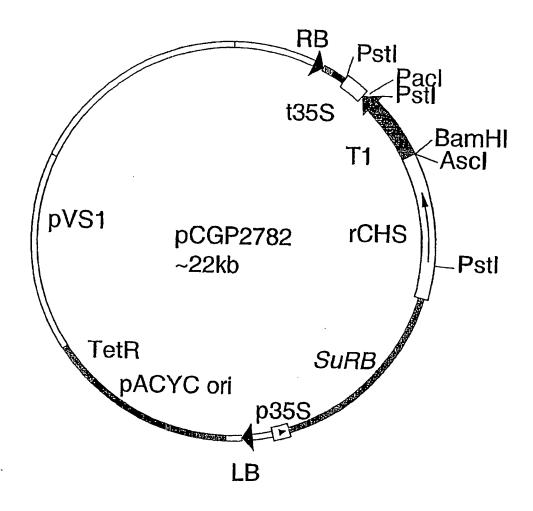
Figure 25



Replicon: pCGP2757 Sall (blunt)/Asp718 ~18.6kb vector

Insert: ~2.7kb <u>HindIII</u> (blunt)/<u>Asp</u>718 fragment from pCGP1116

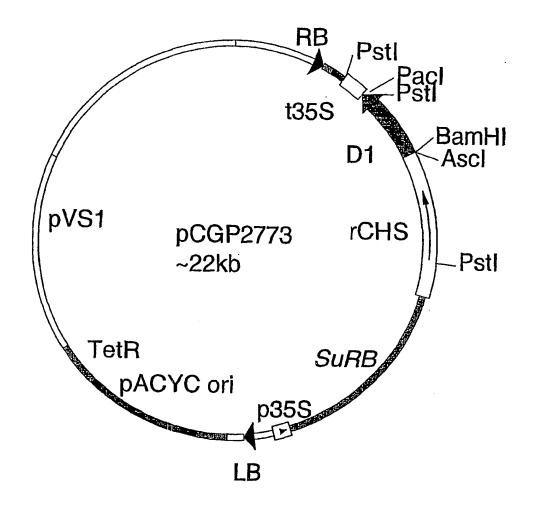
Figure 26



Replicon: pCGP3255 AscI/PacI ~21.3kb vector

Insert: ~0.7kb <u>AscI/PacI T1 PCR</u> product using the primers visproF1 and visproR1 and pCGP2921 as template

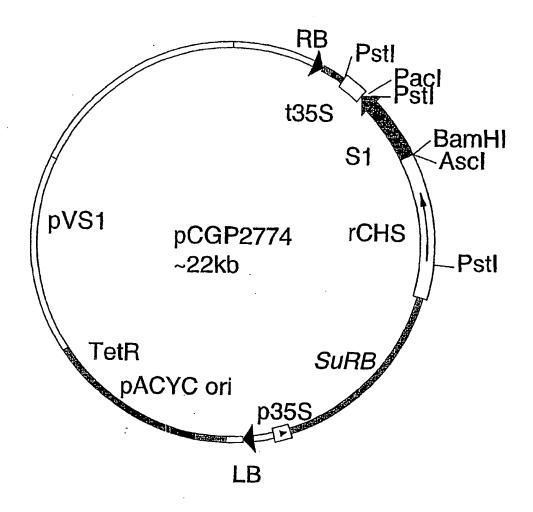
Figure 27



Replicon: pCGP3255 AscI/PacI ~21.3kb vector

Insert: ~0.7kb <u>AscI/PacI</u> D1 PCR product using the primers visproF1 and visproR1 and pCGP2919 as template

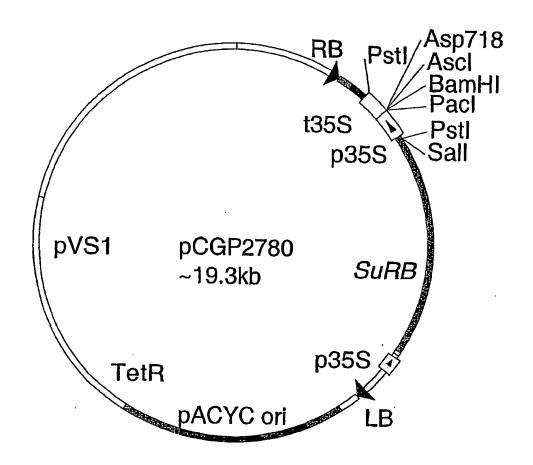
Figure 28



Replicon: pCGP3255 Ascl/Pacl ~21.3kb vector

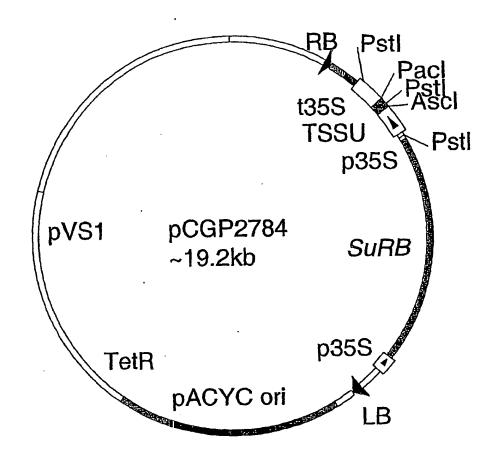
Insert: ~0.7kb <u>Ascl/Pacl</u> S1 PCR product using the primers visproF1 and visproR1 and pCGP2923 as template

Figure 29



Replicon: Religation of <u>Sal</u>I cut pCGP2757 ~19.3kb vector to remove a number of restriction endonuclease recognition sites thereby creating a unique <u>Barn</u>HI site

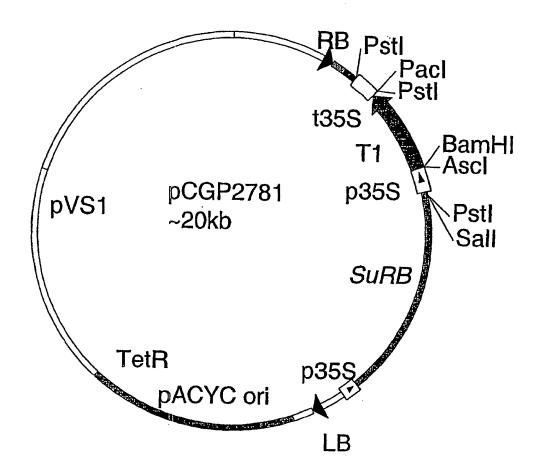
Figure 30



Replicon: pCGP2780 Ascl/BamHI ~19kb vector

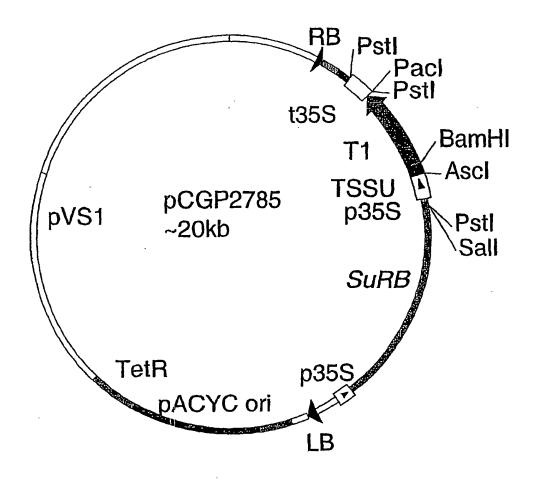
Insert: ~0.2kb <u>Ascl/Bam</u>HI fragment from pCGP2783 containing a chloroplast transit-peptide sequence from tobacco ribulose bisphosphate carboxylase gene

Figure 31



Replicon: Religation of <u>Sal</u>l cut pCGP2772 ~19.3kb vector to remove a number of restriction endonuclease recognition sites thereby creating a unique <u>Bam</u>HI site

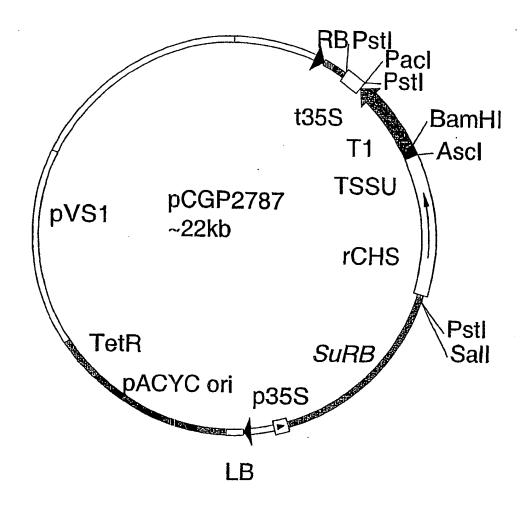
Figure 32



Replicon: pCGP2781 Ascl/BamHI ~20kb vector

Insert: ~0.2kb <u>AscI/Bam</u>HI fragment from pCGP2783 containing a chloroplast transit-peptide sequence from tobacco ribulose bisphosphate carboxylase gene.

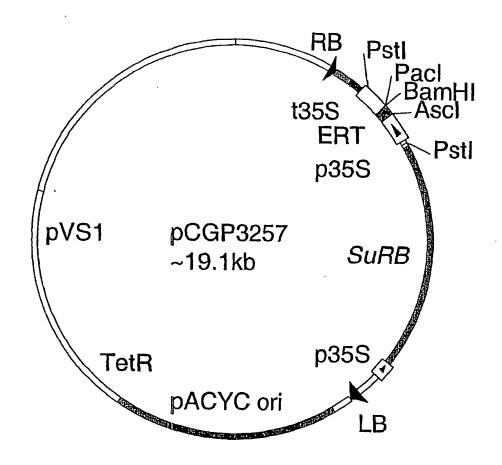
Figure 33



Replicon: pCGP2782 AscI/BamHI ~22kb vector

Insert: ~0.2kb <u>Ascl/Bam</u>HI fragment from pCGP2783 containing a chloroplast transit-peptide sequence from tobacco ribulose bisphosphate carboxylase gene

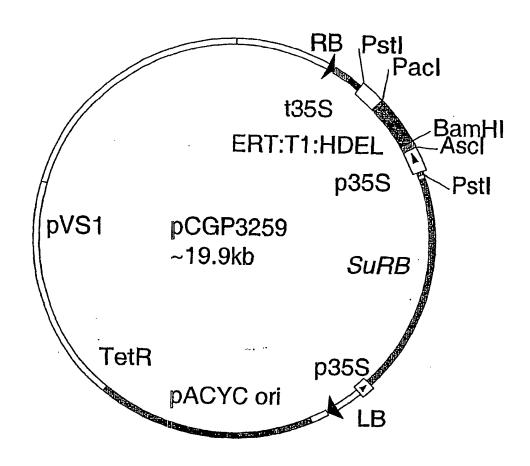
Figure 34



Replicon: pCGP2780 Ascl/BamHl ~19kb vector

Insert: ~0.1kb <u>AscI/Bam</u>HI fragment from pCGP3256 containing an ER-targeting signal sequence from *Arabidopsis* basic chitinase gene

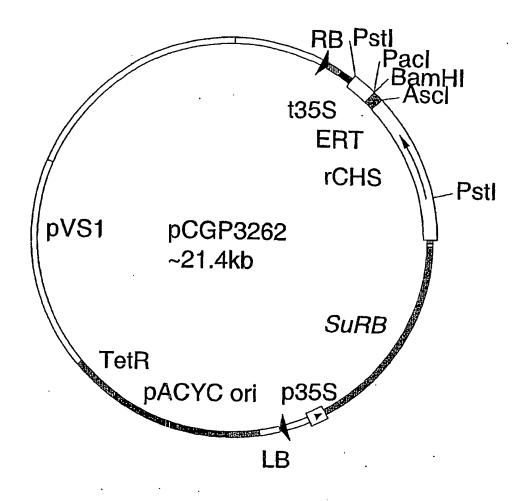
Figure 35



Replicon: pCGP3257 BamHI/PacI ~19.2kb vector

Insert: ~0.7kb <u>BamHI/PacI</u> T1 PCR fragment generated using visproF1 and CPHDELPacR primers and pCGP2779 as template

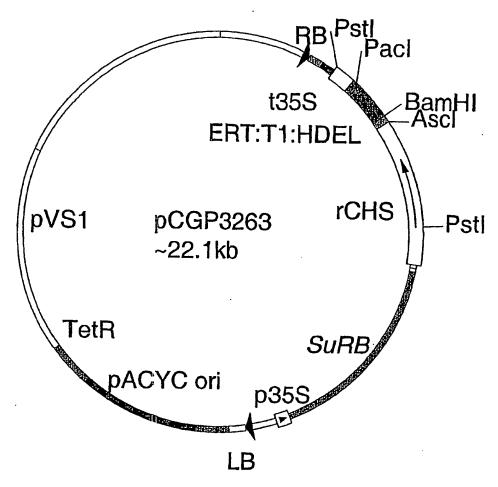
Figure 36



Replicon: pCGP3255 Ascl/BamHI ~21.3kb vector

Insert: ~0.1kb <u>Ascl/Bam</u>HI fragment from pCGP3256 containing an ER-targeting signal sequence from *Arabidopsis* basic chitinase gene

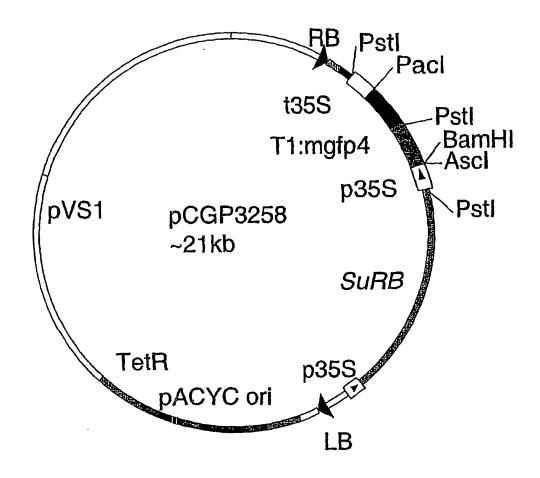
Figure 37



Replicon: pCGP3262 BamHI/PacI ~21.4kb vector

Insert: ~0.7kb <u>BamHI/PacI T1 PCR fragment</u> generated using visproF1 and CPHDELPacR primers and pCGP2779 as template

Figure 38



Replicon: pCGP3257 Ascl/Pacl ~19.4kb vector

Insert: ~1.4kb <u>Asc</u>I/<u>Pac</u>I fragment containing fusion of T1 and mgfp4. T1 PCR product generated using visproF1new and visproRI primers and pCGP2779 as template. mgfp4 PCR product generated using mGFP4-PacIR and Pst-mGFP4F as primers and pBIN35Smgfp4ER as template. PCR products digested with <u>Pst</u>I and ligated together prior to ligation with vector.

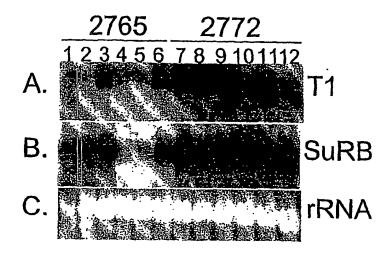


Figure 40

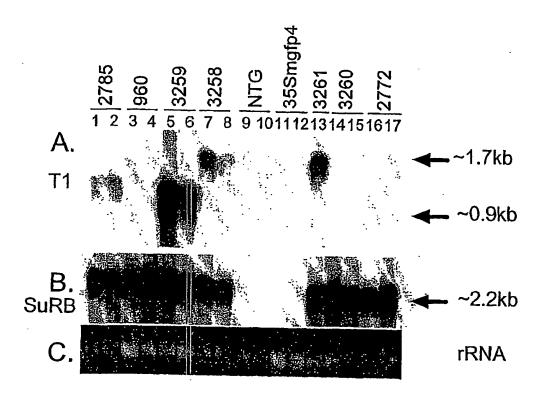


Figure 41

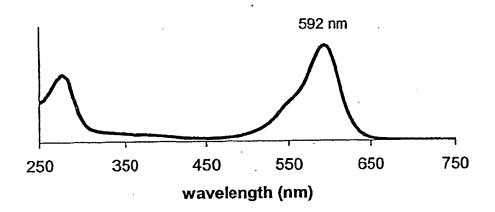


Figure 42(a)

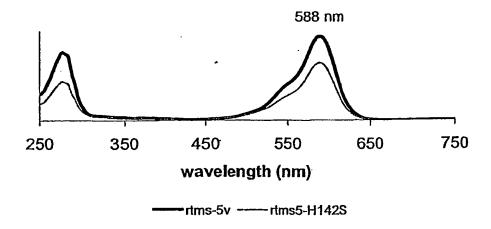


Figure 42(b)

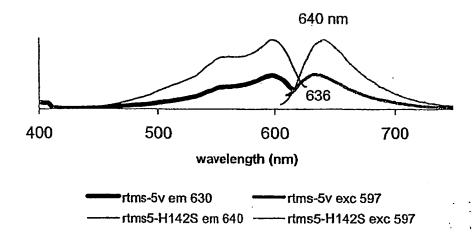


Figure 42(c)

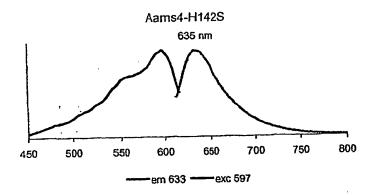


Figure 43(a)

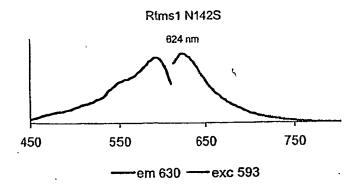
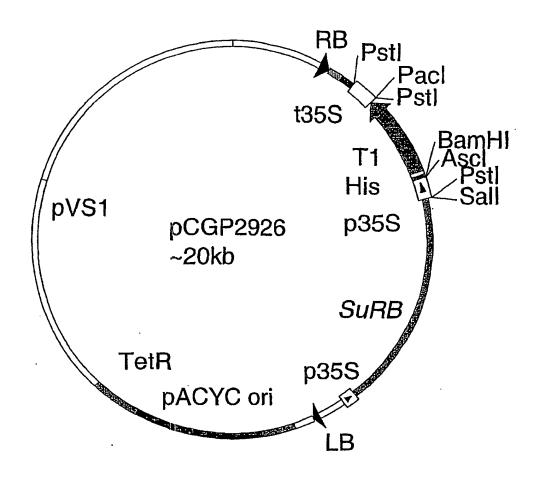


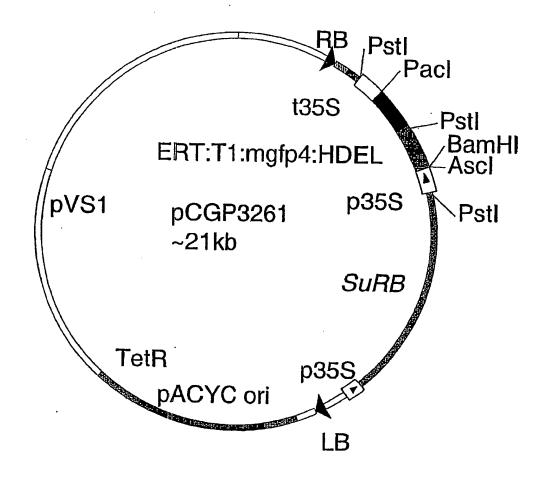
Figure 43(b)



Replicon: pCGP2781 Ascl/Pacl ~20kb vector

Insert: ~0.1kb <u>AscI/PacI</u> fragment containing RBS-TICS and RGSHHHHHH epitope

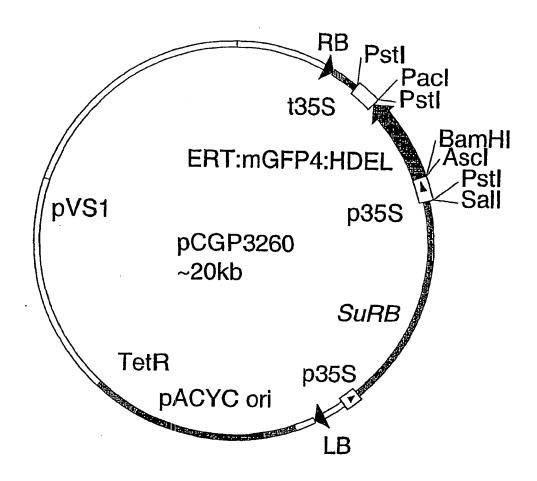
Figure 44



Replicon: pCGP3257 BamHI/PacI ~19.2kb vector

Insert: ~1.4kb <u>BamHI/PacI</u> fragment containing fusion of T1 and mgfp4 amplified from pCGP3258 using primers which incorporated ER targeting (ERT) and retention (HDEL) sequences.

Figure 45



Replicon: pCGP2780 <u>BamHI/PacI</u> (blunt) ~19kb vector

Insert: ~0.7kb <u>Sac</u>I(blunt)/BamHI mGFP4 insert from pBIN35Smgfp4ER which includes ERT and HDEL sequences for ER targeting and retention.

Figure 46

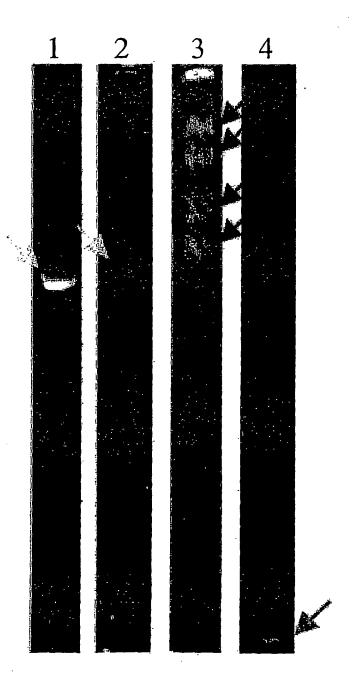


Figure 47

- <110> Nufarm Limited
  - The University of Queensland
- <120> Cell visual characteristic-modifying sequences
- <130> 2505563/EJH
- <140> International
- <141> 2002-03-01
- <150> US 60/273,227
- <151> 2001-03-02
- <150> AU PR3874
- <151> 2001-03-21
- <150> US 60/329,816
- <151> 2001-10-15
- <160> 271
- <170> PatentIn version 3.0
- <210> 1
- <211> 30
- <212> DNA
- <213> oligonucleotide
- <400> 1

tccgtta	atcg ctaaacagat gacctacaaa	30 .
<210>	2	
<211>	21	
<212>	AND	
<213> ·	oligonucleotide	
<400> ggcgaco	2 caca ggtttgcgtg t	21
<210>	3	
<211>	21	
<212>	DNA	
<213>	oligonucleotide	
<400> atgagto	3 gtga tcgctacaca a	21
<210>	4	
<211>	21	
<212>	DNA	
<213>	oligonucloetide	
<400> tttgtg	4 cctt gatttgactc t	21
<210>	5	
<211>	5	
<212>	PRT	
<213>	polypeptide	
<400>	5	
Ser Va. 1	l Ile Ala Lys 5	

2/234

PCT/GB02/00928

WO 02/070703

PCT/GB02/00928

3/234

<210> 6

<211> 6

<212> PRT

<213> polypeptide

<400> 6

Met Ser Val Ile Ala Thr 1 5

<210> 7

<211> 5

<212> PRT

<213> polypeptide

<400> 7

Ser Gly Ile Ala Thr 1 5

<210> 8

<211> 5

<212> PRT

<213> polypeptide

<400> 8

Ser Val Ile Val Thr

<210> 9

<211> 5

<212> PRT

<213> polypeptide

<400> 9

Ser Val Ser Ala Thr 1 5

4/234

<210> 10

<211> 16

<212> PRT

<213> polypeptide

<400> 10

<210> 11

<211> 16

<212> PRT

<213> polypeptide

<400> 11

<210> 12

<211> 16

<212> PRT

<213> polypeptide

<400> 12

Ser Val Ile Ala Thr Gln Val Thr Tyr Lys Val Tyr Met Ser Gly Thr 1  $\phantom{-}5\phantom{+}\phantom{+}\phantom{+}\phantom{+}10\phantom{+}\phantom{+}\phantom{+}\phantom{+}$ 

<210> 13

<211> 16

<212> PRT

<213> polypeptide

<400> 13

Ser Gly Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr 1  $\phantom{\bigg|}$  5  $\phantom{\bigg|}$  10  $\phantom{\bigg|}$  15

5/234

<210> 14

<211> 16

<212> PRT

<213> polypeptide

<400> 14

Ser Val Ile Val Thr Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr 1 5 10 15

<210> 15

<211> 16

<212> PRT

<213> polypeptide

<400> 15

Ser Val Ser Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr 1 5 10 15

<210> 16

<211> 16

<212> PRT

<213> polypeptide

<400> 16

Ser Val Ile Ala Lys Gln Met Thr Tyr Lys Val Asn Met Ser Gly Thr 1  $\phantom{\bigg|}$  5  $\phantom{\bigg|}$  10  $\phantom{\bigg|}$  15

<210> 17

<211> 16

<212> PRT

<213> polypeptide

<400> 17

Ser Val Ile Ala Lys Gln Met Thr Tyr Lys Val Tyr Met Ser Asp Thr 1 5 10 15

```
<210> 18
```

<220>

<220>

<220>

<223> 
$$X = any amino acid except M$$

<400> 18

<210> 19

#### PCT/GB02/00928 WO 02/070703 7/234

<220>													
<221> CDS													
<222> (1)(660)													
<pre>&lt;400&gt; 19 tcc gtt atc gct aaa cag atg acc tac aag gtt tat atg tca gac acg Ser Val Ile Ala Lys Gln Met Thr Tyr Lys Val Tyr Met Ser Asp Thr 1 5 10 15</pre>	48												
gtc aat gga cac tac ttt gag gtt gaa ggc gat gga aaa gga aag cct Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys Pro 20 25 30	96												
tac gag ggg gag cag acg gta aag ctc act gtc acc aag ggc gga cct Tyr Glu Gly Glu Gln Thr Val Lys Leu Thr Val Thr Lys Gly Gly Pro 35 40 45	144												
ctg cca ttt gct tgg gat att cta tca cca cag agt cag tac gga agc Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Ser Gln Tyr Gly Ser 50 55 60	192												
ata cca ttc acc aag tac cct gaa gac atc cct gac tat gta aag cag Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys Gln 65 70 75 80	240												
tca ttc cct gag gga tat aca tgg gag agg atc atg aac ttc gaa gat Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ile Met Asn Phe Glu Asp 85 90 95	288												
ggt gca gtg tgt act gtc agc aat gat tcc agc atc caa ggt aac tgt Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys 100 105 110	336												
ttc atc tac aat gtc aag ttc tct ggt ttg aac ttt cct ccc aat gga Phe Ile Tyr Asn Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly 115 120 125	384												
cct gtt atg caa aag aag aca cag ggc tgg gaa ccc aac act ggg cgt Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Gly Arg 130 135 140	432												
ctc ttt gca cga gat gga atg ctg ata gga aac aac ttt atg gct ctg Leu Phe Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu 145 150 160	480												
aag ttg gaa gga ggt ggt cat tat ttg tgt gaa ttc aaa tct act tac Lys Leu Glu Gly Gly Gly His Tyr Leu Cys Glu Phe Lys Ser Thr Tyr 165 170 175	528												
aag gca aag aag cct gtg atg atg cca ggg tat cac tat gtt gac cgc Lys Ala Lys Lys Pro Val Met Met Pro Gly Tyr His Tyr Val Asp Arg 180 185 190	576												
aag ttg gat gta acc aat cac aac aag gat tac act tcc gtt gag cag Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu Gln 195 200 205	624												

8/234

tgt gaa att tcc att gca cgc aaa cct gtg gtc gcc
Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala
210
215

<210> 20

<211> 220

<212> PRT

<213> Acropora aspera

<400> 20

Ser Val Ile Ala Lys Gln Met Thr Tyr Lys Val Tyr Met Ser Asp Thr 1 5 10 15

Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys Pro 20 25 30

Tyr Glu Gly Glu Gln Thr Val Lys Leu Thr Val Thr Lys Gly Gly Pro 35 40 45

Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Ser Gln Tyr Gly Ser 50 55 60

Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys Gln 65 70 75 80

Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ile Met Asn Phe Glu Asp 85 90 95

Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys 100 105 110

Phe Ile Tyr Asn Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly 115 120 125

Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Gly Arg 130 . 135 140

Leu Phe Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu 145 150 155 160

Lys Leu Glu Gly Gly His Tyr Leu Cys Glu Phe Lys Ser Thr Tyr 165 170 175

# WO 02/070703 PCT/GB02/00928

9/234

	Lys Pro 180	Val Met	Met Pro		His Ty	yr Val 190	Asp	Arg	
Lys Leu Asp 195	Val Thr	Asn His	Asn Lys 200	Asp Tyı		er Val )5	Glu	Gln	
Cys Glu Ile 210	Ser Ile	Ala Arg 215	Lys Pro	Val Val	Ala 220				
<210> 21									
<211> 660									
<212> DNA									
<213> Acrop	pora aspe	ra							
<220>									
<221> CDS									
<222> (1).	. (660)								
<400> 21 tcc gtt atc Ser Val Ile 1									48
tcc gtt atc Ser Val Ile	Ala Lys 5 cac tac	Gln Met ttt gag	Thr Tyr	Lys Val	Tyr Me	et Ser aa gga	Gly 15 aag	Thr	48 96
tcc gtt atc Ser Val Ile 1 gtc aat gga	Ala Lys 5 cac tac His Tyr 20 gag cag	Gln Met ttt gag Phe Glu acg gta	Thr Tyr gtc gaa Val Glu 25	Lys Val 10 ggc gat Gly Asp	Tyr Me	et Ser  aa gga ys Gly 30  ag ggc ys Gly	Gly 15 aag Lys gga	Thr cct Pro	
tcc gtt atc Ser Val Ile 1 gtc aat gga Val Asn Gly tac gag ggg Tyr Glu Gly	Ala Lys 5  cac tac His Tyr 20  gag cag Glu Gln  gct tgg	Gln Met ttt gag Phe Glu acg gta Thr Val gat att	Thr Tyr gtc gaa Val Glu 25 agg ctc Arg Leu 40 tta tca	Lys Value 10 10 10 10 10 10 10 10 10 10 10 10 10	gga aa Gly Ly acc aa Thr Ly 45	aa gga ys Gly 30 ag ggc ys Gly 5	Gly 15 aag Lys gga Gly	Thr cct Pro cct Pro	96
tcc gtt atc Ser Val Ile 1  gtc aat gga Val Asn Gly  tac gag ggg Tyr Glu Gly 35  ctg cca ttt Leu Pro Phe	Ala Lys 5  cac tac His Tyr 20  gag cag Glu Gln  gct tgg Ala Trp  acc aag Thr Lys	Gln Met  ttt gag Phe Glu  acg gta Thr Val  gat att Asp Ile 55  tac cct	Thr Tyr  gtc gaa Val Glu 25 agg ctc Arg Leu 40 tta tca Leu Ser	Lys Value 10 ggc gat Gly Asp act gtc Thr Value cca cac cac Pro Glr	gga aa ag	aa gga ys Gly 30 ag ggc ys Gly 5 ag tac ln Tyr	Gly 15 aag Lys gga Gly gga Gly	Thr  cct Pro  cct Pro  agc Ser	96 144
tcc gtt atc Ser Val Ile 1  gtc aat gga Val Asn Gly  tac gag ggg Tyr Glu Gly 35  ctg cca ttt Leu Pro Phe 50 Phe	Ala Lys 5  cac tac His Tyr 20  gag cag Glu Gln  gct tgg Ala Trp  acc aag Thr Lys  gag gga	Gln Met  ttt gag Phe Glu  acg gta Thr Val  gat att Asp Ile 55  tac cct Tyr Pro 70  tat aca	Thr Tyr  gtc gaa Val Glu 25  agg ctc Arg Leu 40  tta tca Leu Ser  gaa gac Glu Asp	Lys Value 10 ggc gat Gly Asp act gtc Thr Value Cca Cac Cac Cac Cac Cac Cac Cac Cac Cac	gga aa	aa gga ys Gly 30 ag ggc ys Gly 5 ag tac in Tyr at gta yr Val	Gly 15 aag Lys gga Gly gga Gly aag Lys	Thr  cct Pro  cct Pro  agc Ser  cag Gln 80 gat	96 144 192

PCT/GB02/00928

10/234 tto ato tac cat gto aag tto tot ggt ttg aac ttt cot coc aat gga 384 Phe Ile Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly 115 120 cct gtt atg cag aag aca cag ggc tgg gaa ccc aac act gag cgt 432 Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu Arg ctc tta gca cga gat gga atg ctg ata gga aac aac ttt atg gct ctg 480 Leu Leu Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu 150 aag tta gaa gga ggt ggt cac tat ttg tgt gaa ttc aaa tct act tac 528 Lys Leu Glu Gly Gly Gly His Tyr Leu Cys Glu Phe Lys Ser Thr Tyr aag gca agg aag cct gtg aag atg cca ggg tat cac tat gtt gac cgc 576 Lys Ala Arg Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Asp Arg 180 185 aaa ctg gat gta acc aat cac aac aag gat tac act tcc gtt gag cag 624 Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu Gln 200 cgt gaa att tcc att gca cgc aaa cct gtg gtc gcc 660 Arg Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 215 <210> 22 <211> 220 <212> PRT

<213> Acropora aspera

<400> 22

Ser Val Ile Ala Lys Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr

Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys Pro

Tyr Glu Gly Glu Gln Thr Val Arg Leu Thr Val Thr Lys Gly Gly Pro 35

Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Ser Gln Tyr Gly Ser 50 55

Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys Gln 70

#### WO 02/070703 PCT/GB02/00928

### 11/234

Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ile Met Asn Phe Glu Asp 85 90 95

Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys 100 105 110

Phe Ile Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly 115 120 125

Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu Arg 130 135 140

Leu Leu Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu 145 150 155 . 160

Lys Leu Glu Gly Gly Gly His Tyr Leu Cys Glu Phe Lys Ser Thr Tyr 165 170 175

Lys Ala Arg Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Asp Arg 180 185 190

Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu Gln 195 200 205

Arg Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 210 215 220

<210> 23

<211> 663

<212> DNA

<213> Acropora aspera

<220>

<221> CDS

<222> (1)..(663)

<400> 23

tcc gtt atc gct aaa Cag atg acc tac aaa gtt tat atg tca ggc acg Ser Val Ile Ala Lys Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr 1 5 10 15

48

## WO 02/070703 PCT/GB02/00928 12/234

									ggc Gly							96
									act Thr							144
									cca Pro							192
				_			_	-	atc Ile		_		_	_	_	240
									agg Arg 90							288
	_		-		-	_		_	tcc Ser	_					_	336
				_	_				ttg Leu							384
									tgg Trp							432
									gga Gly							480
									tgt Cys 170							528
-	-	_	-				_		ggg Gly				-	-	_	576
	-	_	-					_	gat Asp				-			624
									gtg Val			tga				663

<210> 24

<211> 220

<212> PRT

<213> Acropora aspera

<400> 24

Ser Val Ile Ala Lys Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr 1 5 10 . 15

Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys Pro 20 25 30

Tyr Glu Gly Glu Gln Thr Val Arg Leu Thr Val Thr Lys Gly Gly Pro
35 40 45

Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Ser Gln Tyr Gly Ser 50 55 60

Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys Gln 65 70 75 80

Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ile Met Asn Phe Glu Asp 85 90 95

Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys 100 105 110

Phe Ile Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly 115 120 125

Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu Arg 130 135 140

Leu Phe Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu 145 150 155 160

Lys Leu Glu Gly Gly His Tyr Leu Cys Glu Phe Lys Ser Thr Tyr 165 170 175

Lys Ala Lys Lys Pro Val Arg Met Pro Gly Tyr His Tyr Val Asp Arg 180 185 190

Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu Gln 195 200 205

Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 210 215

<210	>	25														
<211	>	660														
<212	!>	DNA														
<213	3> ;	Acan	thas	tria	sp.											
<220	>															
<221	.> (	CDS														
<222	:>	(1).	. (66	0)												
<400		25														
tcc Ser	gtt Val	atc Ile	gct Ala	Lys	cag Gln	atg Met	acc Thr	tac Tyr	aag Lys	gtt Val	tat Tyr	atg Met	tca Ser	ggc Gly	acg Thr	48
1 .				5					10					15		
gtc Val	aat Asn	gga Gly	His	tac Tyr	ttt Phe	gag Glu	gtc Val	Glu	ggc Gly	gat Asp	gga Gly	aaa Lys	gga Gly	aag Lys	cct Pro	96
			20					25					30			
tac Tyr	gag Glu	Gly ggg	gag Glu	cag Gln	acg Thr	gta Val	aag Lys	ctc Leu	act Thr	gtc Val	acc Thr	aag Lys	ggc Gly	gga Gly	cct Pro	144
		35					40					45				
ctg Leu	Pro	ttt Phe	gct Ala	tgg Trp	gat Asp	Ile	tta Leu	tca Ser	cca Pro	cgg Arg	tgt Cys	cag Gln	tac Tyr	gga Gly	aac Asn	192
	50					55					60					
ata Ile	cca Pro	ttc Phe	acc Thr	aag Lys	Tyr	cct Pro	gaa Glu	gac Asp	gtc Val	cct Pro	gac Asp	tat Tyr	gta Val	aag Lys	cag Gln	240
65					70					75					80	
tca Ser	ttc Phe	ccg Pro	gag Glu	gga Gly	ttt Phe	aca Thr	tgg Trp	gag Glu	agg Arg	atc Ile	atg Met	aac Asn	ttt Phe	gaa Glu	gat Asp	288
				85					90					95		
ggt Gly .																336
			100					105					110			
ttc Phe	acc Thr	tac Tyr	cat His	gtc Val	aag Lys	ttc Phe	tct Ser	ggt Gly	ttg Leu	aac Asn	ttt Phe	cct Pro	ccc Pro	aat Asn	gga Gly	384
		115					120					125				
cct ( Pro																432
	130					135					140					
ctc :		-				_							_	-	_	480
145					150					155					160	

	PCT/GB02/00928
15/234	
	15/234

•		***	J () 2, ()	, , , , , ,						15/	234							,
					ggc Gly 165													528
	aag Lys	gca Ala	aag Lys	aag Lys 180	cct Pro	gtg Val	aag Lys	atg Met	cca Pro 185	G1y ggg	tat Tyr	cat His	tat Tyr	gtt Val 190	gac Asp	cgc Arg		576
					acc Thr													624
					att Ile												,	660
	<210	)> 2	26															
	<211	.> 2	220															
	<212	?> 1	PRT															
	<213	l> 1	Acant	thast	cria	sp.												
	<400	)> 2	26															
	Ser 1	Val	Ile	Ala	Lys 5	Gln	Met	Thr	Tyr	Lys 10	Val	Tyr	Met	Ser	Gly 15	Thr		
	Val	Asn	Gly	His 20	Tyr	Phe	Glu	Val	Gl u 25	Gly	Asp	Gly	Lys	Gly 30	Lys	Pro		
	Tyr	Glu	Gly 35	Glu	Gln	Thr	Val	Lys 40	Leu	Thr	Val	Thr	Lys 45	Gly	Gly	Pro		
	Leu	Pro 50	Phe	Ala	Trp	Asp	Ile 55	Leu	Ser	Pro	Arg	Cys 60	Gln	Tyr	Gly	Asn		
	Ile 65	Pro	Phe	Thr	Lys	Tyr 70	Pro	Glu	Asp	Val	Pro 75	Asp	Tyr	Val	Lys	Gln 80		
	Ser	Phe	Pro	Glu	Gly 85	Phe	Thr	Trp	Glu	Arg 90	Ile	Met	Asn	Phe	Glu .95	Asp		
	Gly	Ala	Val	Cys <sub>.</sub> 100	Thr	Val	Ser	Asn	Asp 105	Ser	Ser	Ile	Gln	Gly 110	Asn	Cys		
	Phe	Thr	Tyr 115	His	Val	Lys	Phe	Ser 120	Gly	Leu	Asn	Phe	Pro 125	Pro	Asn	Gly		

Pro Val Met Gln Lys Lys Thr Arg Gly Trp Glu Pro His Ser Glu Arg 130 135 140	
Leu Phe Ala Arg Gly Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu 145 150 . 155 160	
Lys Leu Glu Gly Gly His Tyr Leu Cys Gly Phe Lys Thr Thr Tyr 165 170 175	
Lys Ala Lys Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Asp Arg 180 185 190	
Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Ile Ser Val Glu Gln 195 200 205	
Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 210 215 220	
<210> 27	
<211> 660	
<212> DNA	
<213> Acanthastria sp.	-
<220>	
<221> CDS	
<222> (1)(660)	
<400> 27	
tcc gtt atc gct aaa cag atg acc tac aaa gtt tat atg tca ggc acg  Ser Val Ile Ala Lys Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr  1 10 15	8
gtc aat gga cac tac ttt gag gtc gaa ggc gat gga aaa gga aag cct 9 Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys Pro 20 25 30	6
tac gag ggg gag cag acg gta agg ctg act gtc acc aag ggc gga cct Tyr Glu Gly Glu Gln Thr Val Arg Leu Thr Val Thr Lys Gly Gly Pro 35 40 45	4

17/234

ata Ile 65	cca Pro	ttc Phe	acc Thr	aag Lys	tac Tyr 70	cct Pro	gaa Glu	gac Asp	atc Ile	cct Pro 75	gac Asp	tat Tyr	gta Val	aag Lys	cag Gln 80	240
tca Ser	ttc Phe	ccg Pro	gag Glu	gga Gly 85	tat Tyr	aca Thr	tgg Trp	gag Glu	agg Arg 90	atc Ile	atg Met	aac Asn	ttt Phe	gaa Glu 95	gat Asp	288
		gtg Val														336
		tac Tyr 115														384
		atg Met														432
		gca Ala														480
_		gaa Glu						_	_	_						528
		agg Arg														576
		gat Asp 195														624
_	-	att Ile			_	_				_	_					660

<210> 28

<211> 220

<212> PRT

<213> Acanthastria sp.

<400> 28

Ser Val Ile Ala Lys Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr 1  $\phantom{\bigg|}$  5  $\phantom{\bigg|}$  10  $\phantom{\bigg|}$  15

Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys Pro 20 25 30

18/234

Tyr Glu Gly Glu Gln Thr Val Arg Leu Thr Val Thr Lys Gly Gly Pro 35 40 45

Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Ser Gln Tyr Gly Ser 50 60

Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys Gln 65 70 75 80

Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ile Met Asn Phe Glu Asp 85 90 95

Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys 100 105 110

Phe Ile Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly 115 120 125

Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu Arg 130 135 140

Leu Ser Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu 145 150 155 160

Lys Leu Glu Gly Gly His Tyr Leu Cys Glu Phe Lys Ser Thr Tyr 165 170 175

Lys Ala Arg Lys Pro Val Lys Met Pro Gly Tyr His Cys Val Asp Arg 180 185 190

Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu Gln 195 200 205

Arg Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 210 215 220

<210> 29

<211> 660

<212> DNA

<213> Acanthastria sp.

<220>

# 19/234

<221> CDS

<222> (1)..(660)

	0> 2 gtt Val															48
gtc Val	aat Asn	gga Gly	cac His 20	tac Tyr	ttt Phe	gag Glu	gtc Val	gaa Glu 25	ggc Gly	gat Asp	gga Gly	aaa Lys	gga Gly 30	aag Lys	cct Pro	96
	gag Glu															144
	cca Pro 50															192
	cca Pro			_			-	_			_		_	_	_	240
	ttc Phe															288
	gca Ala		_		_	_		-		_					_	336
	acc Thr			-	_				_							384
	gtg Val 130															432
	ttt Phe	_				_	-						-	_	_	480
	tta Leu															528
_	gca Ala	-	_			_	_						_	_	-	576
	ctg Leu	_	_					_	_				-		-	624

20/234

tgt gaa atc tcc att gca cgc aaa cct gtg gtc gcc Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 210 215 220 660

<210> 30

<211> 220

<212> PRT

<213> Acanthastria sp.

<400> 30

Ser Val Ile Ala Lys Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr 1 5 10 15

Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys Pro 20 25 30

Tyr Glu Gly Glu Gln Thr Val Lys Leu Thr Val Thr Lys Gly Gly Pro 35 40 45

Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Cys Gln Tyr Gly Ser 50 55 60

Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys Gln 65 70 75 80

Ser Phe Pro Glu Gly Phe Thr Trp Glu Arg Ile Met Asn Phe Glu Asp 85 90 95

Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys 100 105 110

Phe Thr Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly 115 120 125

Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro His Ser Glu Arg 130 135 140

Leu Phe Ala Arg Gly Gly Met Leu Ile Gly Asn Asn Phe Met Ala Pro 145 150 150 160

Lys Leu Glu Gly Gly Gly His Tyr Leu Cys Glu Phe Lys Thr Thr Tyr 165 170 175

21/234

Lys Ala Lys Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Asp Arg
180 185 190

Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu Gln 195 200 205

Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 210 215 220

<210> 31

<211> 660

<212> DNA

<213> Caulastrea ap.

<220>

<221> CDS

<222> (1)..(660)

100

<400> 31 tcc gtt atc gct aaa cag atg acc tac aaa gtt tat atg tca ggc acg 48 Ser Val Ile Ala Lys Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr gtc aat gga cac tac ttt gag gtc gaa ggc gat gga aaa gga aag cct 96 Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys Pro 20 tac gag ggg gag cag acg gta agg ctg act gtc acc aag ggc gga cct 144 Tyr Glu Gly Glu Gln Thr Val Arg Leu Thr Val Thr Lys Gly Gly Pro 40 ctg cca ttt gct tgg gat att tta tca cca cag tca cag tac gga agc 192 Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Ser Gln Tyr Gly Ser 50 ata cca ttc acc aag tac cct gaa gac atc cct gac tat gta aag cag 240 Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys Gln 70 tca ttc ccg gag gga tat aca tgg gag agg atc atq aac ttt gaa gat 288 Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ile Met Asn Phe Glu Asp ggt gca gtg tgt act gtc agc aat gat tcc agc atc caa ggc aac tgt 336

Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys

105

WO 02/070703	PCT/GB02/00928

WO 02/070703 PC 22/234														PCT/C	GB02/0			
								•		22	/234							
t i	tc a ne I	atc [le	tac Tyr 115	cat His	gtc Val	aag Lys	ttc Phe	tct Ser 120	ggt Gly	ttg Leu	aac Asn	ttt Phe	cct Pro 125	ccc Pro	aat Asn	gga Gly		384
P:	ro V	gtt Val 130	atg Met	cag Gln	aag Lys	aag Lys	aca Thr 135	cag Gln	ggc Gly	tgg Trp	gaa Glu	ccc Pro 140	aac Asn	act Thr	gag Glu	cgt Arg		432
$\mathbf{L}_{\mathbf{c}}$	tc t eu S 45	ct Ser	gca Ala	cga Arg	gat Asp	gga Gly 150	atg Met	ctg Leu	ata Ile	gga Gly	aac Asn 155	aac Asn	ttt Phe	atg Met	gct Ala	ctg Leu 160		480
aa Ly	ag t ys I	ta Leu	gaa Glu	gga Gly	ggt Gly 165	ggt Gly	cac His	tat Tyr	ttg Leu	tgt Cys 170	gaa Glu	ttc Phe	aaa Lys	tct Ser	act Thr 175	tac Tyr		528
aa Ly	ag g ys A	lca Ala	agg Arg	aag Lys 180	cct Pro	gtg Val	aag Lys	atg Met	cca Pro 185	ggg Gly	tat Tyr	cac His	tat Tyr	gtt Val 190	gac Asp	cgc Arg		576
aa Ly	aa c ys I	etg Seu	gat Asp 195	gta Val	acc Thr	aat Asn	cac His	aac Asn 200	aag Lys	gat Asp	tac Tyr	act Thr	tcc Ser 205	gtt Val	gag Glu	cag Gln		624
	cg G										gtc Val							660
<2	210>	> 3	32															
<2	211>	> 2	20															
<2	212>	> F	PRT															
<2	213>	· С	aula	stre	ea ap	· .												
< 4	100>	• 3	32															
Se 1	er V	al	Ile	Ala	Lys 5	Gln	Met	Thr	Tyr	Lys 10	Val	Tyr	Met	Ser	Gly 15	Thr		
Vá	al A	sn	Gly	His 20	Tyr	Phe	Glu	Val	Glu 25	Gly	Asp	Gly	Lys	Gly 30	Lys	Pro		
T	r G	Slu	Gly 35	Glu	Gln	Thr	Val	Arg 40	Leu	Thr	Val	Thr	Lys 45	Gly	Gly	Pro		
Le		ro 50	Phe	Ala	Trp	Asp	Ile 55	Leu	Ser	Pro	Gln	Ser 60	Gln	Tyr	Gly	Ser		

Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys Gln 65 70 75 80

### 23/234

Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ile Met Asn Phe Glu Asp 85 90 95

Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys 100 105 110

Phe Ile Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly 115 120 125

Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu Arg 130 135 140

Leu Ser Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu 145 150 155 160

Lys Leu Glu Gly Gly His Tyr Leu Cys Glu Phe Lys Ser Thr Tyr 165 170 175

Lys Ala Arg Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Asp Arg 180 185 190

Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu Gln 195 200 205

Arg Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 210 215 220

<210> 33

<211> 660

<212> DNA

<213> Caulastrea ap.

<220>

<221> CDS

<222> (1)..(660)

<220>

<221> misc feature

<222> (72)..(72)

<223> n = any nucleotide

<220>

<221> misc\_feature

<222> (133)..(133)

<223> n = any nucleotide

<220>

<221> misc\_feature

<222> (166)..(166)

<223> n = any nucleotide

<220>

<221> misc\_feature

<222> (193)..(193)

<223> n = any nucleotide

<220>

<221> misc\_feature

<222> (486)..(486)

<223> n = any nucleotide

<220>

<221> misc\_feature

<222> (495)..(495)

## 25/234

<223> n = any nucleotide

<220>

<221> misc\_feature

<222> (501)..(501)

<223> n = any nucleotide

<400	)> 3	33															
					cag Gln											4 8	3
gtc Val	aat Asn	gga Gly	cac His 20	tac Tyr	ttt Phe	gag Glu	gtn Val	gaa Glu 25	ggc Gly	gat Asp	gga Gly	aaa Lys	gga Gly 30	aag Lys	cct Pro	96	5
				_	acg Thr	_	_			-		_				144	1
					gat Asp											192	2
					tac Tyr 70		_	_			_		-	_	_	240	)
					tat Tyr											288	}
	-		_		gtc Val	-		_		-					_	336	5
				-	aag Lys				_							384	1
					aag Lys											432	2
					gga Gly 150											480	)

### 26/234

aag Lys	ttn Xaa	gaa Glu	gga Gly	ggn Gly 165	ggt Gly	can Thr	tat Tyr	ttg Leu	tgt Cys 170	gaa Glu	ttc Phe	aaa Lys	tct Ser	act Thr 175	tac Tyr	528
aag Lys	gca Ala	aag Lys	aag Lys 180	cct Pro	gtg Val	atg Met	atg Met	cca Pro 185	ggg Gly	tat Tyr	cac His	tat Tyr	gtt Val 190	gac Asp	cgc Arg	576
	ttg Leu															624
	gaa Glu 210															660

<210> 34

<211> 220

<212> PRT

<213> Caulastrea ap.

<220>

<221> misc\_feature

<222> (72)..(72)

<223> n = any nucleotide

<220>

<221> misc\_feature

<222> (133)..(133)

<223> n = any nucleotide

<220>

<221> misc\_feature

<222> (166)..(166)

<223> n = any nucleotide

<220>

<221> misc\_feature

<222> (193)..(193)

<223> n = any nucleotide

<220>

<221> misc\_feature

<222> (486)..(486)

<223> n = any nucleotide

<220>

<221> misc\_feature

<222> (495)..(495)

<223> n = any nucleotide

<220>

<221> misc\_feature

<222> (501)..(501)

<223> n = any nucleotide

<400> 34

Ser Val Ile Ala Lys Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr 1 5 10 15

Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys Pro 20 25 30

Tyr Glu Gly Glu Gln Thr Val Lys Leu Thr Val Thr Xaa Gly Gly Pro  $35 \hspace{1cm} 40 \hspace{1cm} 45$ 

Leu Pro Phe Ala Trp Asp Ile Xaa Ser Pro Gln Ser Gln Tyr Gly Ser 50 55 60

Xaa Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys Gln 65 70 75 80

Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ile Met Asn Phe Glu Asp 85 90 95

Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys . 100 105 110

Phe Ile Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly 115 120 125

Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro His Ser Glu Arg 130 135 140

Leu Phe Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu 145 150 155 160

Lys Xaa Glu Gly Gly Gly Thr Tyr Leu Cys Glu Phe Lys Ser Thr Tyr 165 170 175

Lys Ala Lys Lys Pro Val Met Met Pro Gly Tyr His Tyr Val Asp Arg 180 185 190

Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu Gln 195 200 205

Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 210 215 220

<210> 35

<211> 660

<212> DNA

<213> Caulastrea ap.

<220>

<221> CDS

<222> (1)..(660)

<220>

<221> misc\_feature

<222> (637)..(638)

<223> n = any nucleotide

<220>	
<221>	misc_feature
<222>	(640)(640)
<223>	n = any nucleotide

<400	)> :	35															
tcc Ser 1	gtt Val	atc Ile	gct Ala	aaa Lys 5	cag Gln	atg Met	acc Thr	tac Tyr	aag Lys 10	gtt Val	tat Tyr	atg Met	tca Ser	ggc Gly 15	acg Thr		48
gtc Val	aat Asn	gga Gly	cac His 20	tac Tyr	ttt Phe	gag Glu	gtc Val	gaa Glu 25	ggc Gly	gat Asp	gga Gly	aaa Lys	gga Gly 30	aag Lys	cct Pro		96
tac Tyr	gag Glu	ggg Gly 35	gag Glu	cag Gln	acg Thr	gta Val	aag Lys 40	ctc Leu	act Thr	gtc Val	acc Thr	aag Lys 45	ggc Gly	gga Gly	cct Pro		144
ctg Leu	cca Pro 50	ttt Phe	gct Ala	tgg Trp	gat Asp	att Ile 55	tta Leu	tca Ser	cca Pro	cag Gln	tgt Cys 60	cag Gln	tac Tyr	gga Gly	agc Ser	:	192
ata Ile 65	cca Pro	ttc Phe	acc Thr	aag Lys	tac Tyr 70	cct Pro	gaa Glu	gac Asp	atc Ile	cct Pro 75	gac Asp	tat Tyr	gta Val	aag Lys	cag Gln 80	:	240
tca Ser	ttc Phe	ccg Pro	gag Glu	gga Gly 85	ttt Phe	aca Thr	tgg Trp	gag Glu	agg Arg 90	atc Ile	atg Met	aac Asn	ttt Phe	gaa Glu 95	gat Asp	:	288
ggt Gly	gca Ala	gtg Val	tgt Cys 100	act Thr	gtc Val	agc Ser	aat Asn	gat Asp 105	tcc Ser	agc Ser	atc Ile	caa Gln	ggc Gly 110	aac Asn	tgt Cys	;	336
ttc Phe	acc Thr	tac Tyr 115	cat His	gtc Val	aag Lys	ttc Phe	tct Ser 120	ggt Gly	ttg Leu	aac Asn	ttt Phe	cct Pro 125	ccc Pro	aat Asn	Gly ggg	;	384
cct Pro	gtg Val 130	atg Met	cag Gln	aag Lys	aag Lys	aca Thr 135	cag Gln	ggc Gly	tgg Trp	gaa Glu	ccc Pro 140	cac His	tct Ser	gag Glu	cgt Arg	4	432
ctc Leu 145	ttt Phe	gca Ala	cgg Arg	ggt Gly	gga Gly 150	atg Met	ctg Leu	ata Ile	gga Gly	aac Asn 155	aac Asn	ttt Phe	atg Met	gct Ala	ctg Leu 160		480
aag Lys	tta Leu	gaa Glu	ggg Gly	ggc Gly 165	ggt Gly	cac His	tat Tyr	ttg Leu	tgt Cys 170	gaa Glu	ttc Phe	aaa Lys	act Thr	act Thr 175	tac Tyr	į	528

WO 02/070703	PCT/GB02/00928
--------------	----------------

30/234
aag gca aag aag cet gtg aag atg eca ggg tat eat tat gtt tae age 576 Lys Ala Lys Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Tyr Ser 180 185 190
acc att cat gta acc aat cac aac aag gat tac act tcc gtt gag cag Thr Ile His Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu Gln 195 200 205
tgt gaa att tcc nnt nca cgc aaa cct gtg gtc gcc Cys Glu Ile Ser Xaa Xaa Arg Lys Pro Val Val Ala 210 215 220
<210> 36
<211> 220
<212> PRT
<213> Caulastrea ap.
<220>
<221> misc_feature .
<222> (637)(638)
<223> n = any nucleotide
<220>
<221> misc_feature
<222> (640)(640)
<223> n = any nucleotide
<400> 36
Ser Val Ile Ala Lys Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr 1 5 10 15
Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys Pro 20 25 30 ·
Tyr Glu Gly Glu Gln Thr Val Lys Leu Thr Val Thr Lys Gly Gly Pro 35 40 45

Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Cys Gln Tyr Gly Ser 50 60

Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys Gln 65 70 75 80

Ser Phe Pro Glu Gly Phe Thr Trp Glu Arg Ile Met Asn Phe Glu Asp 85 90 95

Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys 100 105 110

Phe Thr Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly 115 120 125

Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro His Ser Glu Arg 130 135 140

Leu Phe Ala Arg Gly Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu 145 150 155 160

Lys Leu Glu Gly Gly His Tyr Leu Cys Glu Phe Lys Thr Thr Tyr 165 170 175

Lys Ala Lys Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Tyr Ser 180 185 190

Thr Ile His Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu Gln 195 200 205

Cys Glu Ile Ser Xaa Xaa Arg Lys Pro Val Val Ala 210 215 220

<210> 37

<211> 660

<212> DNA

<213> Caulastrea ap.

<220>

<221> CDS

<222> (1)..(660)

<220>

- <221> misc\_feature
- <222> (30)..(30)
- <223> n = any nucleotide
- <220>
- <221> misc\_feature
- <222> (32)..(32)
- <223> n = any nucleotide
- <220>
- <221> misc\_feature
- <222> (37)..(37) -
- <223> n = any nucleotide
- <220>
- <221> misc\_feature
- <222> (48)..(48)
- <223> n = any nucleotide
- <220>
- <221> misc\_feature
- <222> (52)..(52)
- <223> n = any nucleotide

WO 02/070703	PCT/GB02/00928
--------------	----------------

•																
< 400	0> :	37														
tcc Ser 1	gtt Val	atc Ile	gct Ala	aaa Lys 5	cag Gln	atg Met	acc Thr	tac Tyr	atn Xaa 10	gnt Xaa	tat Tyr	ntg Xaa	tca Ser	ggc Gly 15	acn Xaa	48
gtc Val	nat Xaa	gga Gly	cac His 20	tac Tyr	ttt Phe	gag Glu	gtc Val	gaa Glu 25	ggc Gly	gat Asp	gga Gly	aaa Lys	gga Gly 30	aag Lys	cct Pro	96
	gag Glu															144
ctg Leu	cca Pro 50	ttt Phe	gct Ala	tgg Trp	gat Asp	att Ile 55	tta Leu	tca Ser	cca Pro	cag Gln	tgt Cys 60	cag Gln	tac Tyr	gga Gly	agc Ser	192
	cca Pro															240
	ttc Phe															288
	gca Ala															336
ttc Phe	acc Thr	tac Tyr 115	cat His	gtc Val	aag Lys	ttc Phe	tct Ser 120	ggt Gly	ttg Leu	aac Asn	ttt Phe	cct Pro 125	ccc Pro	aat Asn	GJ À aaa	384
	gtg Val 130															432
	ttt Phe															480
	tta Leu															528
	gca Ala															576
	ctg Leu	-	-					_	-				-		_	624
	gaa Glu 210															660

<210> 38

- <211> 220
- <212> PRT
- <213> Caulastrea ap.
- <220>
- <221> misc\_feature
- <222> (30)..(30)
- <223> n = any nucleotide
- <220>
- <221> misc\_feature
- <222> (32)..(32)
- <223> n = any nucleotide
- <220>
- <221> misc\_feature
- <222> (37)..(37)
- <223> n = any nucleotide
- <220>
- <221> misc\_feature
- <222> (48)..(48)
- <223> n = any nucleotide
- <220>
- <221> misc\_feature
- <222> (52)..(52)
- <223> n = any nucleotide
- <400> 38

#### WO 02/070703 PCT/GB02/00928 35/234

Ser Val Ile Ala Lys Gln Met Thr Tyr Xaa Xaa Tyr Xaa Ser Gly Xaa

Val Xaa Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys Pro

Tyr Glu Gly Glu Gln Thr Val Lys Leu Thr Val Thr Lys Gly Gly Pro

Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Cys Gln Tyr Gly Ser 55 60

Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys Gln

Ser Phe Pro Glu Gly Phe Thr Trp Glu Arg Ile Met Asn Phe Glu Asp

Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys 100 105

Phe Thr Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly 115

Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro His Ser Glu Arg 130 135

Leu Phe Ala Arg Gly Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu 145 150 155

Lys Leu Glu Gly Gly His Tyr Leu Cys Glu Phe Lys Thr Thr Tyr

Lys Ala Lys Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Asp Arg

Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu Gln 200 195

Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 215

<210> 39

<211> 660

<212>	DNA															
<213>	Gree	n Po	cill	opor	a											
<220>																
<221>	CDS															
<222>	(1).	. (66	0)													
<400>	39															
tcc g	tt atc al Ile	gct Ala	aaa Lys 5	cag Gln	atg Met	acc Thr	tac Tyr	aag Lys 10	gtt Val	tat Tyr	atg Met	tca Ser	ggc Gly 15	acg Thr		48
gtc a Val A	at gga sn Gly	cac His 20	tac Tyr	ttt Phe	gag Glu	gtt Val	gaa Glu 25	ggc Gly	gat Asp	gga Gly	aaa Lys	gga Gly 30	aag Lys	cct Pro		96
tac g Tyr G	ag ggg lu Gly 35	gag Glu	cag Gln	acg Thr	gta Val	aag Lys 40	ctc Leu	act Thr	gtc Val	acc Thr	aag Lys 45	ggc Gly	gga Gly	cct Pro	•	144
ctg co Leu P:	ca ttt ro Phe O	gct Ala	tgg Trp	gat Asp	att Ile 55	cta Leu	tca Ser	cca Pro	cag Gln	agt Ser 60	cag G1n	tac Tyr	gga Gly	agc Ser		192
ata co Ile P: 65	ca ttc ro Phe	acc Thr	aag Lys	tac Tyr 70	cct Pro	gaa Glu	gac Asp	atc Ile	cct Pro 75	gac Asp	tat Tyr	gta Val	aag Lys	cag Gln 80		240
tca ti Ser Pl	tc cct he Pro	gag Glu	gga Gly 85	tat Tyr	aca Thr	tgg Trp	gag Glu	agg Arg 90	atc Ile	atg Met	aac Asn	ttc Phe	gaa Glu 95	gat Asp		288
ggt go Gly A	ca gtg la Val	tgt Cys 100	act Thr	gtc Val	agc Ser	aat Asn	gat Asp 105	tcc Ser	agc Ser	atc Ile	caa Gln	ggt Gly 110	aac Asn	tgt Cys		336
ttc at Phe II	tc tac le Tyr 115	aat Asn	gtc Val	aag Lys	ttc Phe	tct Ser 120	ggt Gly	ttg Leu	aac Asn	ttt Phe	cct Pro 125	ccc Pro	aat Asn	gga Gly		384
Pro Va	tt atg al Met 30	caa Gln	aag Lys	aag Lys	aca Thr 135	cag Gln	ggc Gly	tgg Trp	gaa Glu	ccc Pro 140	aac Asn	act Thr	gag Glu	cgt Arg		432
ctc tt Leu Pl 145	tt gca he Ala	cga Arg	gat Asp	gga G1y 150	atg Met	ctg Leu	ata Ile	gga Gly	aac Asn 155	aac Asn	t <b>tt</b> Phe	atg Met	gct Ala	ctg Leu 160		480
aag tt Lys Le	tg gaa eu Glu	gga Gly	ggt Gly 165	ggt Gly	cat His	tat Tyr	ttg Leu	tgt Cys 170	gaa Glu	ttc Phe	aaa Lys	tct Ser	act Thr 175	tac Tyr	!	528

								•								
aag Lys	gca Ala	aag Lys	aag Lys 180	cct Pro	gtg Val	atg Met	atg Met	cca Pro 185	ggg Gly	tat Tyr	cac His	tat Tyr	gtt Val 190	gac Asp	cgc Arg	576
aaa Lys	ttg Leu	gat Asp 195	gta Val	acc Thr	aat Asn	cac His	aac Asn 200	aag Lys	gat Asp	tac Tyr	act Thr	tcc Ser 205	gtt Val	gag Glu	cag Gln	624
tgt Cys	gaa Glu 210	att Ile	tcc Ser	att Ile	gca Ala	cgc Arg 215	aaa Lys	cct Pro	gtg Val	gtc Val	gcc Ala 220					660
<210	)> -	40														
<21	l> :	220														
<212	2> 1	PRT														
<213	3> (	Green	n Poo	cillo	pora	ā.										
<400	)> 4	40														
Ser 1	Val	Ile	Ala	Lys 5	Gln	Met	Thr	Tyr	Lys 10	Val	Tyr	Met	Ser	Gly 15	Thr	
Val	Asn	Gly	His 20	Tyr	Phe	Glu	Val	Glu 25	Gly	Asp	Gly	Lys	Glу 30	Lys	Pro	
Туг	Glu	Gly 35	Glu	Gln	Thr	Val	Lys 40	Leu	Thr	Val	Thr	Lys 45	Gly	G1y	Pro	
Leu	Pro 50	Phe	Ala	Trp	Asp	Ile 55	Leu	Ser	Pro	Gln	Ser 60	Gln	Tyr	Gly	Ser	
Ile 65	Pro	Phe	Thr	Lys	Tyr 70	Pro	Glu	Asp	Ile	Pro 75	Asp	Tyr	Val	Lys	G1n 80	
Ser	Phe	Pro	Glu	Gly 85	Tyr	Thr	Trp	Glu	Arg 90	Ile	Met	Asn	Phe	Glu 95	Asp	
Gly	Ala	Val	Cys 100	Thr	Val	Ser	Asn	Asp 105	Ser	Ser	Ile	Gln	Gly 110	Asn	Cys	
Phe	Ile	Tyr 115	Asn	Val	Lys	Phe	Ser 120	Gly	Leu	Asn	Phe	Pro 125	Pro	Asn	Gly	
Pro	Va1 130	Met	Gln	Lys	Lys	Thr 135	Gln	Gly	Trp	Glu	Pro 140	Asn	Thr	Glu	Arg	

Leu Phe Ala 145	Arg Asp Gly 150		e Gly Asn As 155	n Phe Met	Ala Leu 160	
Lys Leu Glu	Gly Gly Gly 165	His Tyr Le	ı Cys Glu Ph 170	e Lys Ser	Thr Tyr 175	
Lys Ala Lys	Lys Pro Val 180	Met Met Pro		s Tyr Val 190	Asp Arg	
Lys Leu Asp 195	Val Thr Asn	His Asn Ly: 200	s Asp Tyr Th	r Ser Val 205	Glu Gln	
Cys Glu Ile 210	Ser Ile Ala	Arg Lys Pro 215	o Val Val Al 22			
<210> 41						
<211> 660						
<212> DNA						
<213> Acrop	oora nobilis					
<220>						
<221> CDS						
<222> (1)	(660)					
<400> 41 tcc gtt atc Ser Val Ile						48
gtc aat gga Val Asn Gly						<sub>.</sub> 96
tac gag ggg Tyr Glu Gly 35						144
ctg cca ttt Leu Pro Phe 50				r Gln Tyr		192
ata cca ttc Ile Pro Phe 65						240

WO 02/070703		PCT/GB02/00928
	39/234	

110

tca ttc cct gag gga tat aca tgg gag agg atc atg aac ttc gaa gat 288 Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ile Met Asn Phe Glu Asp 85 ggt gca gtg tgt act gtc agc aat gat tcc agc atc caa ggt aac tgt 336 Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys 100

ttc atc tac aat gtc aag ttc tct ggt ttg aac ttt cct ccc aat gga 384 Phe Ile Tyr Asn Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly 115

105

cct gtt atg caa aag aag aca cag ggc tgg gaa ccc aac act gag cgt 432 Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu Arg 130 135

ctc ttt gca cga gat gga atg ctg ata gga aac aac ttt atg gct ctg 480 Leu Phe Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu 145 150 155

aag ttg gaa gga ggt ggt cat tat ttg tgt gaa ttc aaa tct act tac 528 Lys Leu Glu Gly Gly His Tyr Leu Cys Glu Phe Lys Ser Thr Tyr 165

aag gca aag aag cct gtg atg atg cca ggg tat cac tat gtt gac cgc 576 Lys Ala Lys Lys Pro Val Met Met Pro Gly Tyr His Tyr Val Asp Arg 185

aaa ttg gat gta acc aat cac aac aag gat tac act tcc gtt gag cag 624 Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu Gln 195 200

tgt gaa att tcc att gca cgc aaa cct gtg gtc gcc 660 Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 215

<210> 42

<211> 220

<212> PRT

<213> Acropora nobilis

<400> 42

Ser Val Ile Ala Lys Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr 10

Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys Pro 25 30

Tyr Glu Gly Glu Gln Thr Val Lys Leu Thr Val Thr Lys Gly Gly Pro 35 40

### 40/234

Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Ser Gln Tyr Gly Ser 50 55 60

Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys Gln 65 70 75 80

Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ile Met Asn Phe Glu Asp 85 90 95

Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys 100 105 110

Phe Ile Tyr Asn Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly 115 . 120 . 125

Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu Arg 130 135 140

Leu Phe Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu 145 150 155 160

Lys Leu Glu Gly Gly His Tyr Leu Cys Glu Phe Lys Ser Thr Tyr 165 170 175

Lys Ala Lys Lys Pro Val Met Met Pro Gly Tyr His Tyr Val Asp Arg 180 185 190

Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu Gln 195 200 205

Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 210 215 220

<210> 43

<211> 660

<212> DNA

<213> Acropora nobilis

<220>

<221> CDS

<222> (1)..(660)

<40	0>	43														
tcc Ser 1	gtt Val	atc Ile	gct Ala	aaa Lys 5	Cag Gln	atg Met	acc Thr	tac Tyr	aag Lys 10	gtt Val	tat Tyr	atg Met	tca Ser	ggc Gly 15	acg Thr	48
gtc Val	aat Asn	gga Gly	cac His 20	tac Tyr	ttt Phe	gag Glu	gtc Val	gaa Glu 25	ggc Gly	gat Asp ,	gga Gly	aaa Lys	gga Gly 30	aag Lys	cct Pro	96
tac Tyr	gag Glu	ggg Gly 35	gag Glu	cag Gln	acg Thr	gta Val	aag Lys 40	ctc Leu	act Thr	gtc Val	acc Thr	aag Lys 45	ggc Gly	gga Gly	cct Pro	144
		ttt Phe														192
ata Ile 65	cca Pro	ttc Phe	acc Thr	aag Lys	tac Tyr 70	cct Pro	gaa Glu	gac Asp	atc Ile	cct Pro 75	gac Asp	tat Tyr	gta Val	aag Lys	cag Gln 80	240
		ccg Pro														288
		gtg Val													tgt Cys	336
		tac Tyr 115														384
		atg Met														432
		gca Ala														480
		gaa Glu														528
		aag Lys														576
		gat Asp 195														624
		att Ile													•	660

<210> 44

<211> 220

<212> PRT

<213> Acropora nobilis

<400> 44

Ser Val Ile Ala Lys Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr 1 5 10 15

Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys Pro 20 25 30

Tyr Glu Gly Glu Gln Thr Val Lys Leu Thr Val Thr Lys Gly Gly Pro 35 40 45

Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Cys Gln Tyr Gly Ser 50 55 60

Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys Gln 65 70 75 80

Ser Phe Pro Glu Gly Phe Thr Trp Glu Arg Ile Met Asn Phe Glu Asp 85 90 95

Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys 100 105 110

Phe Thr Tyr His Val Lys Phe Ser Gly Leu Asp Phe Pro Pro Asn Gly 115 120 125

Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro His Ser Glu Arg 130 135 140

Leu Phe Ala Arg Gly Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu 145 150 155 160

Lys Leu Glu Gly Gly His Tyr Leu Cys Glu Phe Lys Thr Thr Tyr 165 170 175

Lys Ala Lys Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Asp Arg 180 185 190

### 43/234

Lys Leu Asp	Val	Thr	Asn	His	Asn	Lys	Asp	Tyr	Thr	Ser	Val	Glu	Gln
195					200					205			

Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 210 215 220

<210> 45

<211> 660

<212> DNA

<213> Acropora nobilis

<220>

<221> CDS

<222> (1)..(660)

< 40	0>	45															
tcc Ser 1	gtt Val	atc Ile	gct Ala	aaa Lys 5	cag Gln	atg Met	acc Thr	tac Tyr	aag Lys 10	gtt Val	tat Tyr	atg Met	tca Ser	ggc Gly 15	acg Thr	•	48
gtc Val	aat Asn	gga Gly	cac His 20	tac Tyr	ttt Phe	gag Glu	gtt Val	gaa Glu 25	ggc Gly	gat Asp	gga Gly	aaa Lys	gga Gly 30	aag Lys	cct Pro	!	96
tac Tyr	gag Glu	ggg Gly 35	gag Glu	cag G1n	acg Thr	gta Val	aag Lys 40	ctc Leu	act Thr	gtc Val	acc Thr	aag Lys 45	ggc Gly	gga Gly	cct Pro	1	44
ctg Leu	cca Pro 50	ttt Phe	gct Ala	tgg Trp	gat Asp	att Ile 55	cta Leu	tca Ser	cca Pro	cag Gln	agt Ser 60	cag Gln	tac Tyr	gga Gly	agc Ser	19	92
ata Ile 65	cca Pro	ttc Phe	acc Thr	aag Lys	tac Tyr 70	cct Pro	gaa Glu	gac Asp	atc Ile	cct Pro 75	gac Asp	tat Tyr	gta Val	aag Lys	cag Gln 80	24	40
tca Ser	ttc Phe	cct Pro	gag Glu	gga Gly 85	tat Tyr	aca Thr	tgg Trp	gag Glu	agg Arg 90	atc Ile	atg Met	aac Asn	ttc Phe	gaa Glu 95	gat Asp	28	88
			tgt Cys 100													33	36
ttc Phe	atc Ile	tac Tyr 115	aat Asn	gtc Val	aag Lys	ttc Phe	tct Ser 120	ggt Gly	ttg Leu	aac Asn	ttt Phe	cct Pro 125	ccc Pro	aat Asn	gga Gly	38	84

WO 02/070703		PCT/GB02/00928
	1.1/0.2.1	

cct gtt atg caa aag aag aca cag ggc tgg gaa ccc aac act gag cgt 432
Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu Arg 130 135 140

ctc ttt gca cga gat gga atg ctg ata gga aac aac ttt atg gct ctg 480
Leu Phe Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu 145 150 155 160

aag ttg gaa gga ggt ggt cat tat ttg tgt gaa ttc aaa tct act tac 528
Lys Leu Glu Gly Gly Gly His Tyr Leu Cys Glu Phe Lys Ser Thr Tyr

aag gca aag aag cct gtg atg atg cca ggg tat cac tat gtt gac cgc
Lys Ala Lys Lys Pro Val Met Met Pro Gly Tyr His Tyr Val Asp Arg
180 185 190

aaa ttg gat gta acc aat cac aac aag gat tac act tcc gtt gag cag
Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu Gln
195 200 205

tgt gaa att tcc att gca cgc aaa cct gtg gtc gcc
Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala
210 215 220

<210> 46

<211> 220

<212> PRT

<213> Acropora nobilis

<400> 46

Ser Val Ile Ala Lys Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr 1 5 10 15

Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys Pro 20 25 30

Tyr Glu Gly Glu Gln Thr Val Lys Leu Thr Val Thr Lys Gly Gly Pro 35 40 45

Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Ser Gln Tyr Gly Ser 50 55 60

Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys Gln 65 70 75 80

Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ile Met Asn Phe Glu Asp 85 90 95

45/234

Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys 100 105 110

Phe Ile Tyr Asn Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly 115 120 125

Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu Arg 130 135 140

Leu Phe Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu 145 150 155 160

Lys Leu Glu Gly Gly Gly His Tyr Leu Cys Glu Phe Lys Ser Thr Tyr 165 170 175

Lys Ala Lys Lys Pro Val Met Met Pro Gly Tyr His Tyr Val Asp Arg 180 185 190

Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu Gln 195 200 205

Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 210 215 220

<210> 47

<211> 660

<212> DNA

<213> Acropora nobilis

<220>

<221> CDS

<222> (1)..(660)

<400> 47

tcc gtt atc gct aaa cag atg acc tac aag gtt tat atg tca ggc acg Ser Val Ile Ala Lys Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr 1 5 10 15

gtc aat gga cac tac ttt gag gtt gaa ggc gat gga aaa gga aag cct 96 Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys Pro 20 25 30

WO 02/070703		PCT/GB02/00928
	46/234	

										40/	234						
ta Ty	ıc (	gag Glu	ggg Gly 3 <b>5</b>	gag Glu	cag Gln	acg Thr	gta Val	aag Lys 40	ctc Leu	act Thr	gtc Val	acc Thr	gag Glu 45	Gly Ggc	gga Gly	cct Pro	144
ct Le	eu I	cca Pro 50	ttt Phe	gct Ala	tgg Trp	gat Asp	att Ile 55	cta Leu	tca Ser	cca Pro	cag Gln	agt Ser 60	cag Gln	tac Tyr	gga Gly	agc Ser	192
at 11 65	.e I	cca Pro	ttc Phe	acc Thr	aag Lys	tac Tyr 70	cct Pro	gaa Glu	gac Asp	atc Ile	cct Pro 75	gac Asp	tat Tyr	gta Val	aag Lys	cag Gln 80	240
to Se	er I	ttc Phe	cct Pro	gag Glu	gga Gly 85	tat Tyr	aca Thr	tgg Trp	gag Glu	agg Arg 90	atc Ile	atg Met	aac Asn	ttc Phe	gaa Glu 95	gat Asp	288
G]	jt <u>(</u> .y <i>P</i>	gca Ala	gtg Val	tgt Cys 100	act Thr	gtc Val	agc Ser	aat Asn	gat Asp 105	tcc Ser	agc Ser	atc Ile	caa Gln	ggt Gly 110	aac Asn	tgt Cys	336
t t Ph	c a ie ]	atc Ile	tac Tyr 115	aat Asn	gtc Val	aag Lys	ttc Phe	tct Ser 120	ggt Gly	ttg Leu	aac Asn	ttt Phe	cct Pro 125	ccc Pro	aat Asn	gga Gly	384
Pr	7 o	gtt Val 130	atg Met	caa Gln	aag Lys	aag Lys	aca Thr 135	cag Gln	ggc Gly	tgg Trp	gaa Glu	ccc Pro 140	aac Asn	act Thr	gag Glu	cgt Arg	432
ct Le 14	u E	ttt Phe	gca Ala	cga Arg	gat Asp	gga Gly 150	atg Met	ctg Leu	ata Ile	gga Gly	aac Asn 155	aac Asn	ttt Phe	atg Met	gct Ala	ctg Leu 160	480
aa Ly	g t s I	tg Leu	gaa Glu	gga Gly	ggt Gly 165	ggt Gly	cat His	tat Tyr	ttg Leu	tgt Cys 170	gaa Glu	ttc Phe	aaa Lys	tct Ser	act Thr 175	tac Tyr	528
aa Ly	gg	gca Ala	aag Lys	aag Lys 180	cct Pro	gtg Val	atg Met	atg Met	cca Pro 185	GJÀ āāā	tat Tyr	cac His	tat Tyr	gtt Val 190	gac Asp	cgc Arg	576
aa Ly	a t 's I	tg Leu	gat Asp 195	gta Val	acc Thr	aat Asn	cac His	aac Asn 200	aag Lys	gat Asp	tac Tyr	act Thr	tcc Ser 205	gtt Val	gag Glu	cag Gln	624
	s G			tcc Ser													660

<210> 48

<211> 220

<212> PRT

<213> Acropora nobilis

<400> 48

47/234

Ser Val Ile Ala Lys Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr 1  $\phantom{-}5\phantom{+}\phantom{+}\phantom{+}\phantom{+}10\phantom{+}\phantom{+}\phantom{+}\phantom{+}$ 

Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys Pro 20 25 30

Tyr Glu Gly Glu Gln Thr Val Lys Leu Thr Val Thr Glu Gly Gly Pro
35 40 45

Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Ser Gln Tyr Gly Ser 50 60

Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys Gln 65 70 75 80

Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ile Met Asn Phe Glu Asp 85 90 95

Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys 100 105 110

Phe Ile Tyr Asn Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly 115 120 125

Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu Arg 130 135 140

Leu Phe Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu 145 150 155 160

Lys Leu Glu Gly Gly His Tyr Leu Cys Glu Phe Lys Ser Thr Tyr 165 170 175

Lys Ala Lys Lys Pro Val Met Met Pro Gly Tyr His Tyr Val Asp Arg 180 185 190

Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu Gln 195 200 205

Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 210 215 220

<210> 49

<211> 660

<21	2>	DNA														
<21	3>	Millepora sp. (Hydrozoan)														
<22	<220>															
<22	1>	CDS														
<22	<222> (1)(660)															
	<400> 49 tcc gtt atc gct aaa cag atg acc tac aag gtt tat atg tca ggc acg 48															
Ser 1	Val	Ile	Ala	Lys 5	Gln	Met	Thr	Tyr	Lys 10	Val	Tyr	Met	Ser	Gly 15	Thr	48
g <b>t</b> c Val	aat Asn	gga Gly	cac His 20	tac Tyr	ttt Phe	gag Glu	gtt Val	gaa Glu 25	ggc Gly	gat Asp	gga Gly	aaa Lys	gga Gly 30	aag Lys	cct Pro	96
tac Tyr	gag Glu	999 Gly 35	gag Glu	cag Gln	acg Thr	gta Val	aag Lys 40	ctc Leu	act Thr	gtc Val	acc Thr	aag Lys 45	ggc Gly	gga Gly	cct Pro	144
ctg Leu	cca Pro 50	ttt Phe	gct Ala	tgg Trp	gat Asp	att Ile 55	cta Leu	tca Ser	cca Pro	cag Gln	agt Ser 60	cag Gln	tac Tyr	gga Gly	agc Ser	192
ata Ile 65	cca Pro	ttc Phe	acc Thr	aag Lys	tac Tyr 70	cct Pro	gaa Glu	gac Asp	atc Ile	cct Pro 75	gac Asp	tat Tyr	gta Val	aag Lys	cag Gln 80	240
tca Ser	ttc Phe	cct Pro	gag Glu	gga Gly 85	tat Tyr	aca Thr	tgg Trp	gag Glu	agg Arg 90	atc Ile	atg Met	aac Asn	ttt Phe	gaa Glu 95	gat Asp	288
ggt Gly	gca Ala	gtg Val	tgt Cys 100	act Thr	gtc Val	agc Ser	aat Asn	gat Asp 105	tcc Ser	agc Ser	atc Ile	caa Gln	ggc Gly 110	aac Asn	tgt Cys	336
ttc Phe	atc Ile	tac Tyr 115	cat His	gtc Val	aag Lys	ttc Phe	tct Ser 120	ggt Gly	ttg Leu	aac Asn	ttt Phe	cct Pro 125	ccc Pro	aat Asn	gga Gly	384
cct Pro	gtt Val 130	atg Met	cag Gln	aag Lys	aag Lys	aca Thr 135	cag Gln	ggc Gly	tgg Trp	gaa Glu	ccc Pro 140	aac Asn	act Thr	gag Glu	cgt Arg	432
ctc Leu 145	ttt Phe	gca Ala	cga Arg	gat Asp	gga Gly 150	atg Met	ctg Leu	ata Ile	gga Gly	aac Asn 155	aac Asn	ttt Phe	atg Met	gct Ala	ctg Leu 160	480
aag Lys	tta Leu	gaa Glu	gga Gly	ggt Gly 165	ggt Gly	cac His	tat Tyr	tta Leu	tgt Cys 170	gaa Glu	ttc Phe	aaa Lys	tct Ser	act Thr 175	tac Tyr	528

#### PCT/GB02/00928

WO 02/070703 49/234 aag gca agg aag cct gtg aag atg cca ggg tat cac tat gtt gac cgc 576 Lys Ala Arg Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Asp Arg aaa ctg gat gta acc aat cac aac aag gat tac act tcc gtt gag cag 624 Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu Gln cgt gaa att tcc att gca cgc aaa cct gtg gtc gcc 660 Arg Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 215 <210> 50 <211> 220 <212> PRT <213> Millepora sp. (Hydrozoan) <400> 50 Ser Val Ile Ala Lys Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys Pro Tyr Glu Gly Glu Gln Thr Val Lys Leu Thr Val Thr Lys Gly Gly Pro

Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Ser Gln Tyr Gly Ser

Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys Gln 70

Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ile Met Asn Phe Glu Asp 85 90

Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys 100 105

Phe Ile Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly 115 120

Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu Arg 130 135

50/234

Leu Phe Ala 145	Arg Asp Gly		e Gly Asn As 155	on Phe Met	Ala Leu 160
Lys Leu Glu	Gly Gly Gly 165	His Tyr Le	u Cys Glu Ph 170	ne Lys Ser	Thr Tyr 175
Lys Ala Arg	Lys Pro Val 180	Lys Met Pr 18		is Tyr Val 190	Asp Arg
Lys Leu Asp 195	Val Thr Asr	His Asn Ly 200	s Asp Tyr Th	or Ser Val 205	Glu Gln
Arg Glu Ile 210	Ser Ile Ala	Arg Lys Pr 215	o Val Val Al 22		
<210> 51					
<211> 660					
<212> DNA					
<213> Millo	epora sp. (E	ydrozoan)			
<220>					
<221> CDS					
<222> (1).	. (660)				
<400> 51	act ass can	ata aca ta	att to	·+ -+- +	~~~ ^^
tcc gtt atc Ser Val Ile 1	Ala Lys Gln 5	Met Thr Ty	Lys Val Ty	r Met Ser	ggc acg 48 Gly Thr 15
gtc aat gga Val Asn Gly	cac tac ttt His Tyr Phe 20	gag gtc gaa Glu Val Glu 25	a ggc gat gg ı Gly Asp Gl	ga aaa gga .y Lys Gly 30	aag cct 96. Lys Pro
tac gag ggg Tyr Glu Gly 35					
ctg cca ttt Leu Pro Phe 50				er Gln Tyr	
ata cca ttc Ile Pro Phe 65	acc aag tac Thr Lys Tyr 70	cct gaa gae Pro Glu Asp	e atc cct ga o Ile Pro As 75	c tat gta p Tyr Val	aag cag 240 Lys Gln 80

WO 02/070703	PCT/GB02/00928
WO 02/070703	PCT/GB02/009

										•						
				gga Gly 85												288
ggt Gly	gca Ala	gtg Val	tgt Cys 100	act Thr	gtc Val	agc Ser	aat Asn	gat Asp 105	tcc Ser	agc Ser	atc Ile	caa Gln	ggc Gly 110	aac Asn	tgt Cys	336
				gtc Val												384
				aag Lys												432
				gat Asp												480
				ggt Gly 165												528
				cct Pro												576
	_	_	-	acc Thr				_	_				_		-	624
-	-			att Ile	-	_				-	-					660
<210	)> '	52														
<211		220														
<212		PRT														
<213			epora	a sp.	. (H <u>y</u>	ydroz	zoan)	<b>)</b> .								
<400	)> 5	52														
Ser 1	Val	Ile	Ala	Lys 5	Gln	Met	Thr	Tyr	Lys 10	Val	Tyr	Met	Ser	Gly 15	Thr	
Val	Asn	Gly	His 20	Tyr	Phe	Glu	<b>V</b> al	Glu 25	Gly	Asp	Gly	Lys	Gly 30	Lys	Pro	

Tyr Glu Gly Glu Gln Thr Val Arg Leu Thr Val Thr Lys Gly Gly Pro 35 40 45

#### 52/234

Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Ser Gln Tyr Gly Ser 50 55

Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys Gln 65 70 75 80

Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ile Met Asn Phe Glu Asp 85 90 95

Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys 100 105 110

Phe Ile Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly 115 120 125

Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu Arg 130 135 140

Leu Phe Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu 145 150 155 160

Lys Leu Glu Gly Gly His Tyr Leu Cys Glu Phe Lys Ser Thr Tyr 165 170 175

Lys Ala Arg Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Asp Arg 180 185 190

Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu Gln 195 200 205

Arg Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 210 215 220

<210> 53

<211> 660

<212> DNA

<213> Millepora sp. (hydrozoan)

<220>

<221> CDS

<222> (1)..(660)

< 40	-	53															
tcc Ser 1	gtt Val	atc Ile	gct Ala	aaa Lys 5	cag Gln	atg Met	acc Thr	tac Tyr	aaa Lys 10	gtt Val	tat Tyr	atg Met	tca Ser	ggc Gly 15	acg Thr		48
gtc Val	aat Asn	gga Gly	cac His 20	tac Tyr	ttt Phe	gag Glu	gtc Val	gaa Glu 25	ggc Gly	gat Asp	gga Gly	aaa Lys	gga Gly 30	aag Lys	cct Pro		96
tac Tyr	gag Glu	ggg Gly 35	gag Glu	cag Gln	acg Thr	gta Val	agg Arg 40	ctg Leu	act Thr	gtc Val	acc Thr	aag Lys 45	ggc Gly	gga Gly	cct Pro		144
ctg Leu	cca Pro 50	ttt Phe	gct Ala	tgg Trp	gat Asp	att Ile 55	tta Leu	tca Ser	cca Pro	cag Gln	tca Ser 60	càg Gln	tac Tyr	gga Gly	agc Ser		192
ata Ile 65	cca Pro	ttc Phe	acc Thr	aag Lys	tac Tyr 70	cct Pro	gaa Glu	gac Asp	atc Ile	cct Pro 75	gac Asp	tat Tyr	gta <sup>.</sup> Val	aag Lys	cag Gln 80		240
tca Ser	ttc Phe	ccg Pro	gag Glu	gga Gly 85	tat Tyr	aca Thr	tgg Trp	gag Glu	agg Arg 90	atc Ile	atg Met	aac Asn	ttt Phe	gaa Glu 95	gat Asp		288
	gca Ala																336
	atc Ile																384
cct Pro	gtt Val 130	atg Met	cag Gln	aag Lys	aag Lys	aca Thr 135	cag Gln	ggc Gly	tgg Trp	gaa Glu	ccc Pro 140	aac Asn	act Thr	gag Glu	cgt Arg	,	432
	ttt Phe																480
	tta Leu															!	528
	gca Ala															. !	576
	ctg Leu															1	624
_	gaa Glu 210				_	_				-	_					ı	660

<210> 54

<211> 220

<212> PRT

<213> Millepora sp. (hydrozoan)

<400> 54

Ser Val Ile Ala Lys Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr 1  $\phantom{\bigg|}$  5  $\phantom{\bigg|}$  10  $\phantom{\bigg|}$  15

Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys Pro 20 25 30

Tyr Glu Gly Glu Gln Thr Val Arg Leu Thr Val Thr Lys Gly Gly Pro 35 40 45

Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Ser Gln Tyr Gly Ser 50 55 60

Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys Gln65707580

Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ile Met Asn Phe Glu Asp 85 90 95

Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys 100 105 110

Phe Ile Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly 115 120 125

Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu Arg 130 135 140

Leu Phe Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu 145 150 155 160

Lys Leu Glu Gly Gly His Tyr Leu Cys Glu Phe Lys Ser Thr Tyr . 165 170 175

Lys Ala Arg Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Asp Arg 180 185 190

### 55/234

Lys	Leu	Asp	Val	Thr	Asn	His	Asn	Lys	Asp	Tyr	Thr	Ser	Val	Glu	Gln
		195					200					205			

Arg Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 210 215 220

<210> 55

<211> 660

<212> DNA

<213> Porites Murrayensis

<220>

<221> CDS

<222> (1)..(660)

<400	)> !	55										
		atc Ile										48
		gga Gly										96
		ggg Gly 35										144
		ttt Phe										192
		ttc Phe										240
		cct Pro						_		_	_	288
	_	gtg Val	_	_		-	_					336
		tac Tyr 115										384

CT/GB02/00928

	W	O 02/	07070	3					56	6/234						PCT/GB02/0092
cct Pro	gtt Val 130	atg Met	caa Gln	aag Lys	aag Lys	aca Thr 135	cag Gln	ggc Gly	tgg Trp	gaa Glu	ccc Pro 140	aac Asn	act Thr	GJ À āāā	cgt Arg	432
ctc Leu 145	ttt Phe	gca Ala	cga Arg	gat Asp	gga Gly 150	atg Met	ctg Leu	ata Ile	gga Gly	aac Asn 155	aac Asn	ttt Phe	atg Met	gct Ala	ctg Leu 160	480
			gga Gly													528
aag Lys	gca Ala	aag Lys	aag Lys 180	cct Pro	gtg Val	atg Met	atg Met	cca Pro 185	GJ À GGG	tat Tyr	cac His	tat Tyr	gtt Val 190	gac Asp	cgc Arg	576
			gta Val													624
			tcc Ser													660
<210	)> !	56														
<211	l> 2	220														
<212	?> I	PRT														
<213	3> 1	Porit	tes 1	⁄urra	ayens	sis										
<400	)> 5	56														
Ser 1	Val	Ile	Ala	Lys 5	Gln	Met	Thr	Tyr	Lys 10	Val	Tyr	Met	Ser	Asp 15	Thr	
Val	Asn	Gly	His 20	Tyr	Phe	Glu	Val	Glu 25	Gly	Asp	Gly	Lys	Gly 30	Lys	Pro	
Tyr	Glu	Gly 35	Glu	Gln	Thr	Val	Lys 40	Leu	Thr	Val	Thr	Lys 45	Gly	Gly	Pro	
Leu	Pro 50	Phe	Ala	Trp	Asp	Ile 55	Leu	Ser	Pro	Gln	Ser 60	Gln	Tyr	Gly	Ser	

Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys Gln 65 70 75 80

57/234

Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys 100 105 110

Phe Ile Tyr Asn Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly 115 120 125

Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Gly Arg 130 135 140

Leu Phe Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu 145 150 155

Lys Leu Glu Gly Gly Gly His Tyr Leu Cys Glu Phe Lys Ser Thr Tyr 165 170 175

Lys Ala Lys Lys Pro Val Met Met Pro Gly Tyr His Tyr Val Asp Arg 180 185 190

Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu Gln 195 200 205

Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 210 215 220

<210> 57

<211> 660

<212> DNA

<213> Porites Murrayensis

<220>

<221> CDS

<222> (1)..(660)

<400> 57

tcc gtt atc gct aaa cag atg acc tac aag gtt tat atg tca ggc acg
Ser Val Ile Ala Lys Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr
1 10 15

gtc aat gga cac tac ttt gag gtt gaa ggc gat gga aaa gga agg cct 96 Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Arg Pro 20 25 30 .

WO 02/070703	PCT/GB02/00928

tac Tyr	gag Glu	999 Gly 35	gag Glu	cag Gln	acg Thr	gta Val	aag Lys 40	ctc Leu	act Thr	gtc Val	acc Thr	aag Lys 45	ggc Gly	gga Gly	cct Pro	144
			gct Ala													192
ata Ile 65	cca Pro	ttc Phe	acc Thr	aag Lys	tac Tyr 70	cct Pro	gaa Glu	gac Asp	atc Ile	cct Pro 75	gac Asp	tat Tyr	gta Val	aag Lys	cag Gln 80	240
			gag Glu													288
			tgt Cys 100													336
			aat Asn													384
			cag Gln	_	_					-					-	432
		-	cga Arg	-		_	_						_	_	_	480
			gga Gly													528
			aag Lys 180													576
			gta Val													624
-	_	_	tcc Ser	_	-	_				-	_					660

<210> 58

<211> 220

<212> PRT

<213> Porites Murrayensis

#### 59/234

Ser Val Ile Ala Lys Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr 1 5 10 15

Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Arg Pro 20 25 30

Tyr Glu Gly Glu Gln Thr Val Lys Leu Thr Val Thr Lys Gly Gly Pro 35 40 45

Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Ser Gln Tyr Gly Ser 50 60

Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys Gln 65 70 75 80

Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ile Met Lys Phe Glu Asp 85 90 95

Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Met Gln Gly Asn Cys 100 105 110

Phe Ile Tyr Asn Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly 115 120 125

Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu Arg 130 135 140

Leu Tyr Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu 145 150 155 160

Lys Leu Glu Gly Ser Gly His Tyr Thr Cys Glu Phe Lys Ser Thr Tyr 165 170 175

Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu Gln 195 200 205

Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 210 215 220

<210> 59

<211> 656

# WO 02/070703 PCT/GB02/00928 60/234

00/

<212>	DNA														
<213>	Pori	tes :	Murr	ayens	sis										
<220>															
<221>	CDS														
<222>	(1).	. (50	7)												
<400> tcc gt Ser Va															48
gtc aa Val As	at gga sn Gly	cac His 20	tac Tyr	ttt Phe	gag Glu	gtt Val	gaa Glu 25	ggc	gat Asp	gga Gly	aaa Lys	gga Gly 30	aag Lys	cct Pro	96
tac ga Tyr Gl															144
ctg co Leu Pr 50	o Phe														192
ata co Ile Pr 65	a ttc o Phe	acc Thr	aag Lys	tac Tyr 70	cct Pro	gag Glu	gac Asp	atc Ile	cct Pro 75	gac Asp	tat Tyr	gta Val	aag Lys	cag Gln 80	240
tca tt Ser Ph															288
ggt go Gly Al															336
ttc at Phe Il	_	Asn				_		_							384
cct gt Pro Va 13	1 Met														432
ctt ta Leu Ty 145															480
aag tt Lys Le								tgat	gato	gee a	agg <u>g</u> t	atca	ac		527

61/234

tatgttgacc gcgaattgga tgtaaccaat cacaacaagg attacacttc cgttgagcag 587
tgtgagattt ccatcgcacg caaacctgtg gtcgcctgac gtttttcag agtcaaatca 647
aggcacaaa 656

<210> 60

<211> 169

<212> PRT

<213> Porites Murrayensis

<400> 60

Ser Val Ile Ala Lys Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr 1 5 10 15

Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys Pro  $20 \\ 25 \\ 30$ 

Tyr Glu Gly Glu Gln Thr Val Lys Leu Thr Val Thr Lys Gly Gly Pro 35 40 45

Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Ser Gln Tyr Gly Ser 50 60

Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys Gln 65 70 75 80

Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ile Val Asn Phe Glu Asp 85 90 95

Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys 100 105 110

Phe Ile Tyr Asn Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly 115 120 125

Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu Arg 130 135 140

Leu Tyr Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu 145 150 155 160

Lys Leu Glu Gly Gly Gln Arg Ser Leu 165

WO 02/070703 PCT	/GB02/00928
------------------	-------------

<210	>	61														
<2112	>	693														
<212	>	DNA														
<213	> :	Pori	tes l	Murr	ayen:	sis										
<220	>															
<221	> (	CDS										٠				
<222	>	(1).	. (66	0)												
<400		61														
tcc o	Val	Ile	Ala	Lys	Gln	Met	Thr	Tyr	Lys	gtt Val	Tyr	atg Met	tca Ser	Gly	acg Thr	48
	a a t	aas	020	5	+++	~~~	~+~		10					15	,	0.5
gtc a Val A	Asn	Gly	His 20	Tyr	Phe	Glu	Val	Gln 25	Gly	Asp	gga Gly	Lys	Gly	aag Lys	Pro	96
tac q	nen	aaa		Cau	aca	at a	229		act	ata	200	229	30	~~~		144
Tyr (																144
ctg d	сса		act	taa	gat	att		tca	cca	can	agt		tac	nn a	200	192
Leu E	Pro 50	Phe	Ala	Trp	Asp	Ile 55	Leu	Ser	Pro	G1n	Ser 60	Gln	Tyr	Glý	Ser	132
ata c	cca	ttc	acc	aag	tac		gaa	gac	atc	cct		tat	qta	aag	cag	240
Ile E 65	Pro	Phe	Thr	Lys	Tyr 70	Pro	Ğlu	Āsp	Ile	Pro 75	Åsp	Tyr	Val	Lys	Gln 80	
tca t	ttc	cct	gag	gga	tat	aca	tgg	gag	agg	atc	atg	aac	ttc	aaa	gat	288
Ser E																
ggt g																336
Gly A	Ala	Val	Cys 100	Thr	Val	Ser	Asn	Asp 105	Ser	Ser	Ile	G1n	Gly 110	Asn	Cys	
ttc a																384
Phe 1	тте	Tyr 115	Asn	Val	Lys	Phe	Ser 120	GLy	Leu	Asn	Phe	Pro 125	Pro	Asn	Gly	
cct c	gtt	atg	caa	aag	aag	aca	cag	ggc	tgg	gaa	CCC	aac	act	gag	cgt	432
Pro V	130	met	GTU	ոչ	гу	135	GTII	GIĀ	тф	GIU	140	ASN	mr	GIU	нгg	
ctc t Leu F		_	-	-			_						_	_	_	480
145		71 <b>T</b> G	13± Y	113P	150	- 41	<b>⊅</b> cu	116	OLY	155	UOII	EHG	net	n14	160	

	W	0 02/0	70703	3					63	/234						PCT/GE	302/0
aag Lys	ttg Leu	gaa Glu	gga Gly	ggt Gly 165	ggt Gly	cat His	tat Tyr	ttg Leu	tgt Cys 170	gaa Glu	ttc Phe	aaa Lys	tct Ser	act Thr 175	tac Tyr	!	528
aag Lys	gca Ala	aag Lys	aag Lys 180	cct Pro	gtg Val	atg Met	atg Met	cca Pro 185	GJ À ààà	tat Tyr	cac His	tat Tyr	gtt Val 190	gac Asp	cgc Arg	!	576
aaa Lys	ttg Leu	gat Asp 195	gta Val	acc Thr	aat Asn	cac His	aac Asn 200	aag Lys	gat Asp	tac Tyr	act Thr	tcc Ser 205	gtt Val	gag Glu	cag Gln	(	624
		att Ile										tga	cgtt <sup>.</sup>	ttt		(	670
tca	gagt	caa a	atcaa	aggca	ac aa	aa										(	693
<210	0>	62															
<21	1> :	220															
<212	2> ]	PRT															
<213	3> 1	Pori	tes N	durra	ayens	sis							1				
<400	O> (	62															
Ser 1	Val	Ile	Ala	Lys 5	Gln	Met	Thr	Tyr	Lys 10	Val	Tyr	Met	Ser	Gly 15	Thr		
Val	Asn	Gly	His 20	Tyr	Phe	Glu	Val	Gln 25	Gly	Asp	Gly	Lys	Gly 30	Lys	Pro		
Tyr	Glu	Gly 35	Glu	Gln	Thr	Val	Lys 40	Leu	Thr	Val	Thr	Lys 45	Gly	Gly	Pro		
Leu	Pro 50	Phe	Ala	Trp	Asp	Ile 55	Leu	Ser	Pro	Gln	Ser 60	Gln	Tyr	Gly	Ser		
Ile 65	Pro	Phe	Thr	Lys	Tyr 70	Pro	Glu	Asp	Ile	Pro 75	Asp	Tyr	Val	Lys	Gln 80		
Ser	Phe	Pro	Glu	Gly 85	Tyr	Thr	Trp	Glu	Arg 90	Ile	Met	Asn	Phe	Lys 95	Asp		
Gly	Ala	Val	Cys 100	Thr	Val	Ser	Asn	Asp 105	Ser	Ser	Ile	Gln	Gly 110	Asn	Cys		

Phe Ile Tyr Asn Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly 115 120 125

64/234

130	Met	Gln	Lys	Lys	Thr 135	Gln	Gly	Trp	Glu	Pro 140	Asn	Thr	Glu	Arg	
Leu Phe 145	Ala	Arg	Asp	Gly 150	Val	Leu	Ile	Gly	Asn 155	Asn	Phe	Met	Ala	Leu 160	
Lys Leu	Glu	Gly	Gly 165	Gly	His	Tyr	Leu	Cys 170	Glu	Phe	Lys	Ser	Thr 175	Tyr	
Lys Ala	Lys	Lys 180	Pro	Val	Met	Met	Pro 185	Gly	Tyr	His	Tyr	Val 190	Asp	Arg	
Lys Leu	Asp 195	Val	Thr	Asn	His	Asn 200	Lys	Asp	Tyr	Thr	Ser 205	Val	Glu	Gln	
Cys Glu 210		Ser	Ile	Ala	Arg 215	Lys	Pro	Val	Val	Ala 220					
<210>	63														
<211>	660														
<212>	DNA														
<213>	Platy	gyra	sp.												
<220>															
	CDS														
<221>	CDS (1)	(660	))												
<221> <222> <400>	(1) 63														
<221> <222>	(1) 63 atc	gct	aaa	cag Gln	atg Met	acc Thr	tac Tyr	aag Lys 10	gtt Val	tat Tyr	atg Met	tca Ser	ggc Gly 15	acg Thr	48
<221> <222> <400> tcc gtt Ser Val	(1) 63 atc Ile	gct Ala Cac	aaa Lys 5 tac	Gln ttt	Met gag	Thr gtc	Tyr gaa	Lys 10 ggc	Val gat	Tyr aga	Met aaa	Ser gga	Gly 15 aag	Thr	48 . 96
<221> <222> <400> tcc gtt Ser Val 1 gtc aat	63 atc Ile gga Gly	gct Ala cac His 20	aaa Lys 5 tac Tyr	Gln ttt Phe	Met gag Glu gta	Thr gtc Val	Tyr gaa Glu 25 ctc	Lys 10 ggc Gly act	Val gat Asp gtc	Tyr aga Arg acc	Met aaa Lys aag	gga Gly 30	Gly 15 aag Lys	Thr cct Pro	

WO 02/070703 PCT/GB02/00928 65/234

ata	cca	ttc	acc	aag	tac	cct	gaa	gac	gtc	cct	gac	tat	gta	aaq	cag	:	240
Ile	Pro	Phe	Thr	Lys	Tyr	Pro	Glu	Asp	Val	Pro	Asp	Tyr	Val	Lys	Gln		
65					70					75	_	-		-	80		

											_					
tca	ttc	ccg	gag	gga	LLL	aca	tgg	gag	agg	atc	atg	aac	ttt	gaa	gat	288
Ser	Phe	Pro	Glu	Gly	Phe	Thr	Trp	Glu	Arq	Ile	Met	Asn	Phe	Glu	Asp	
				85					90 ້					95	-	

<210> 64

<211> 220

<212> PRT

<213> Platygyra sp.

<400> 64

Ser Val Ile Ala Lys Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr 1  $\phantom{000}5\phantom{000}$  10  $\phantom{000}15\phantom{000}$ 

Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Arg Lys Gly Lys Pro 20 25 30

66/234

Tyr Glu Gly Glu Gln Thr Val Lys Leu Thr Val Thr Lys Gly Gly Pro
35 40 45

Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Cys Gln Tyr Gly Asn 50 60

Ile Pro Phe Thr Lys Tyr Pro Glu Asp Val Pro Asp Tyr Val Lys Gln 65 70 75 80

Ser Phe Pro Glu Gly Phe Thr Trp Glu Arg Ile Met Asn Phe Glu Asp 85 90 95

Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys 100 105 110

Phe Thr Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly 115 120 125

Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro His Ser Glu Arg 130 135 140

Leu Phe Ala Arg Gly Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu 145 150 155 160

Lys Leu Glu Gly Gly His Tyr Leu Cys Gly Phe Lys Thr Thr Tyr 165 170 175

Lys Ala Lys Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Asp Arg 180 185 190

Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Ile Ser Val Glu Gln
195 200 205

Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 210 215 220

<210> 65

<211> 660

<212> DNA

<213> Platygyra sp.

<220>

67/234

<221> CDS

<222> (1)..(660)

<400 tcc Ser 1	gtt	65 atc Ile	gct Ala	aaa Lys 5	cag Gln	atg Met	acc Thr	tac Tyr	aag Lys 10	gtt Val	tat Tyr	atg Met	tca Ser	ggc Gly 15	acg Thr	48
gtc Val	aat Asn	gga Gly	cac His 20	tac Tyr	ttt Phe	gag Glu	gtc Val	gaa Glu 25	ggc Gly	gat Asp	ej A aaa	aaa Lys	gga Gly 30	aag Lys	cct Pro	96
		ggg Gly 35														144
		ttt Phe														192
		ttc Phe														240
		ccg Pro									_			_	_	288
	_	gtg Val	_		_	_		-		_					_	336
		tac Tyr 115		_	_				_							384
		atg Met														432
		gca Ala														480
_		gaa Glu						Leu	-	Ğly						528
		aag Lys														576
		gat Asp 195														624 ·

tgt gaa att tcc att gca cgc aaa cct gtg gtc gcc
Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala
210 215 220

<210> 66

<211> 220

<212> PRT

<213> Platygyra sp.

<400> 66

Ser Val Ile Ala Lys Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr 1 5 10 15

Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys Pro 20 25 30

Tyr Glu Gly Glu Gln Thr Val Lys Leu Thr Val Thr Lys Gly Gly Pro 35 40 45

Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Cys Gln Tyr Gly Asn 50 55 60

Ile Pro Phe Thr Lys Tyr Pro Glu Asp Val Pro Asp Tyr Val Lys Gln 65 70 75 80

Ser Phe Pro Glu Gly Phe Thr Trp Glu Arg Ile Met Asn Phe Glu Asp  $85 \hspace{1cm} 90 \hspace{1cm} 95 \hspace{1cm} .$ 

Phe Thr Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly 115 120 125

Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro His Ser Glu Arg 130 135 140

Leu Phe Ala Arg Gly Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu 145 150 155 160

Lys Leu Glu Gly Gly Gly His Tyr Leu Cys Gly Phe Lys Thr Thr Tyr 165 170 175

#### 69/234

Lys Ala Lys Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Asp Arg 180 185 190

Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Ile Ser Val Glu Gln 195 200 205

Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 210 215 220

<210> 67

<211> 660

<212> DNA

<213> Platygyra sp.

<220>

<221> CDS

<222> (1)..(660)

<400	)> (	57															
	gtt Val		_		_	_			_	-		_			_	4	8
-	aat Asn						_	_		-				_		9	6
	gag Glu		-	_	_	-	-			_		_				14	4
-	cca Pro 50		_		_					_	-	_				19	2
	cca Pro			_			_	_	_		-		-	-	_	24	0
	ttc Phe															28	8
	gca Ala		_		-	-		-		_					-	33	6

105

70/234

									70/	234						
ttc Phe	acc Thr	tac Tyr 115	cat His	gtc Val	aag Lys	ttc Phe	tct Ser 120	ggt G1y	ttg Leu	aac Asn	ttt Phe	cct Pro 125	ccc Pro	aat Asn	gga Gly	384
cct Pro	gtg Val 130	atg Met	cag Gln	aag Lys	aag Lys	aca Thr 135	cag Gln	ggc Gly	tgg Trp	gaa Glu	ccc Pro 140	cac His	tct Ser	gag Glu	cgt Arg	432
ctc Leu 145	ttt Phe	gca A1a	cgg Arg	ggt Gly	gga Gly 150	atg Met	ctg Leu	ata Ile	gga Gly	aac Asn 155	aac Asn	ttt Phe	atg Met	gct Ala	ctg Leu 160	480
			gga Gly													528
			aag Lys 180													576
			gta Val													624
tgt Cys	gaa Glu 210	act Thr	tcc Ser	att Ile	gca Ala	cgc Arg 215	aaa Lys	cct Pro	gtg Val	gtc Val	gcc Ala 220					660
<210	D> 6	8														
<211	1> 2	220														
<212	2> I	PRT														
<213	3> I	Platy	/gyra	a sp.	•											
<400	)> 6	88														
Ser 1	Val	Ile	Ala	Lys 5	Gln	Met	Thr	Tyr	Lys 10	Val	Tyr	Met	Ser	Gly 15	Thr	
Val	Asn	Gly	His 20	Tyr	Phe	Glu	Val	Glu 25	Gly	Asp	Gly	Lys	G1y 30	Lys	Pro	
Tyr	Glu	Gly 35	Glu	Gln	Thr	Va1	Lys 40	Leu	Thr	Val	Thr	Lys 45	Gly	Gly	Pro	
Leu	Pro 50	Phe	Ala	Trp	Asp	Ile 55	Leu	Ser	Pro	Gln	Cys 60	Gln	Tyr	Gly	Asn	

Ile Pro Phe Thr Lys Tyr Pro Glu Asp Val Pro Asp Tyr Val Lys Gln 65 70 75 80

#### 71/234

Ser Phe Pro Glu Gly Phe Thr Trp Glu Arg Ile Met Asn Phe Glu Asp 85 90 95

Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys 100 105 110

Phe Thr Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly 115 120

Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro His Ser Glu Arg 130 135 140

Leu Phe Ala Arg Gly Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu 145 150 155 160

Lys Leu Glu Gly Gly His Tyr Leu Cys Gly Phe Lys Thr Thr Tyr 165 170 175

Lys Ala Lys Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Asp Arg 180 185 190

Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Ile Ser Val Glu Gln 195 200 205

Cys Glu Thr Ser Ile Ala Arg Lys Pro Val Val Ala 210 215 220

<210> 69

<211> 623

<212> DNA

<213> Platygyra sp.

<220>

<221> CDS

<222> (1)..(621)

<400> 69

tcc gtt atc gct aaa cag atg acc tac aag gtt tat atg tca ggc acg Ser Val Ile Ala Lys Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr 1 5 10 15

48

WO 02/070703		PCT/GB02/00928
	72/234	

gtc Val	aat Asn	gga Gly	cac His 20	tac Tyr	ttt Phe	gag Glu	gtc Val	gaa Glu 25	ggc Gly	gat Asp	gga Gly	aaa Lys	gga Gly 30	aag Lys	cct Pro	96	
tac Tyr	gag Glu	ggg Gly 35	gag Glu	cag Gln	acg Thr	gta Val	aag Lys 40	ctc Leu	act Thr	gtc Val	acc Thr	aag Lys 45	ggc Gly	gga Gly	cct Pro	144	
	cca Pro 50															192	
	cca Pro															240	
	ttc Phe	_									_			_	_	288	
	gca Ala															336	
	acc Thr															384	
	gtg Val 130	_	_	-	-					_					_	432	
	ttt Phe	-				_	_						_	_	-	480	
	ttg Leu									aga Arg 170	-		-			528	
-	ttg Leu		-			_				aca Thr 185						576	
_	ttg Leu	-	_					_		_		_		_	cc	623	

<210> 70

<211> 169

<212> PRT

<213> Platygyra sp.

<400> 70

Ser Val Ile Ala Lys Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr 1  $\phantom{-}5\phantom{+}10\phantom{+}15\phantom{+}15\phantom{+}10\phantom{+}15\phantom{+}15\phantom{+}10\phantom{+}15\phantom{+}10\phantom{+}15\phantom{+}10\phantom{+}15\phantom{+}10\phantom{+}1$ 

Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys Pro 20 25 30

Tyr Glu Gly Glu Gln Thr Val Lys Leu Thr Val Thr Lys Gly Gly Pro 35 40 45

Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Arg Cys Gln Tyr Gly Asn 50 55 60

Ile Pro Phe Thr Lys Tyr Pro Glu Asp Val Pro Asp Tyr Val Lys Gln 65 70 75 80

Ser Phe Pro Glu Gly Phe Thr Trp Glu Arg Ile Met Asn Phe Glu Asp 85 90 95

Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys 100 105 110

Phe Thr Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly 115 120 125

Pro Val Met Gln Lys Lys Thr Arg Gly Trp Glu Pro His Ser Glu Arg 130 135 140

Leu Phe Ala Arg Gly Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu 145 150 155 160

Lys Leu Glu Gly Gly Gln Arg Ser Leu

<210> 71

<211> 13

<212> PRT

<213> Platygyra sp.

<400> 71

Arg Cys Gln Gly Ile Ile Met Leu Thr Ala Asn Trp Met 1 5 10  $\iota$ .

WO 02/070703		PCT/GB02/00928
	74/234	

<210> 72 <211> 23 <212> PR'

<213> Platygyra sp.

<400> 72

Pro Ile Thr Thr Arg Ile Thr Phe Pro Leu Ser Ser Val Lys Phe Pro 1 5 10 15

Leu His Ala Asn Leu Trp Ser 20

<210> 73

<211> 660

<212> DNA

<213> Platygyra sp.

<220>

<221> CDS

<222> (1)..(660)

<400> 73

tc	gtt	atc	gct	aaa	cag	atg	acc	tac	aag	gtt	tat	atg	tca	ggc	acg	48
Sea	Val	Ile	Ala	Lys	Gln	Met	Thr	Tyr	Lys	Val	Tyr	Met	Ser	Glv	Thr	
1				5				•	10		-			15		

gtc aat gga cac tac ttt gag gtt gaa ggc gat gga aaa gga aag cct 96 Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys Pro 20 25 30

tac gag ggg gag cag acg gta aag ctc act gtc acc aag ggc gga cct
Tyr Glu Gly Glu Gln Thr Val Lys Leu Thr Val Thr Lys Gly Gly Pro
35 40 45

ctg cca ttt gct tgg gat att cta tca cca cag agt cag tac gga agc 192 Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Ser Gln Tyr Gly Ser 50 55 60

ata cca ttc acc aag tac cct gaa gac atc cct gac tat gta aag cag

Ile Pro Phe Thr Lys. Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys Gln

75

80

	w	02/0	70703	3						PCT/GB02/009						
									75/	234						
	ttc Phe															288
	gca Ala															336
ttc Phe	atc Ile	tac Tyr 115	aat Asn	gtc Val	aag Lys	ttc Phe	tct Ser 120	ggt Gly	ttg Leu	aac Asn	ttt Phe	cct Pro 125	ccc Pro	aat Asn	gga Gly	384
	gtt Val 130															432
	ttt Phe															480
	ttg Leu															528
_	gca Ala	_	_			_	_						_	_	_	576
	ttg Leu															624
_	gaa Glu 210				-	_				_	_					660

<210> 74

<211> 220

<212> PRT

<213> Platygyra sp.

<400> 74

Ser Val Ile Ala Lys Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr 5 10

Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys Pro

Tyr Glu Gly Glu Gln Thr Val Lys Leu Thr Val Thr Lys Gly Gly Pro 40

# 76/234

Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Ser Gln Tyr Gly Ser 50 55 60

Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys Gln 65 70 75 80

Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ile Met Asn Phe Glu Asp 85 90 95

Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys 100 105 110

Phe Ile Tyr Asn Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly 115 120 125

Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu Arg 130 135 140

Leu Phe Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu 145 150 155 160

Lys Leu Glu Gly Gly His Tyr Leu Cys Glu Phe Lys Ser Thr Tyr 165 170 175

Lys Ala Lys Lys Pro Val Met Met Pro Gly Tyr His Tyr Val Asp Arg 180 185 190

Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu Gln 195 200 205

Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 210 215 220

<210> 75

<211> 660

<212> DNA

<213> Platygyra sp.

<220>

<221> CDS

<222> (1)..(660)

<400>																
tcc gt Ser Va 1	t atc l Ile	gct Ala	aaa Lys 5	cag Gln	atg Met	acc Thr	tac Tyr	aag Lys 10	gtt Val	tat Tyr	atg Met	tca Ser	ggc Gly 15	acg Thr		48
gtc aa Val As	t gga n Gly	cac His 20	tac Tyr	ttt Phe	gag Glu	gtt Val	gaa Glu 25	ggc Gly	gat Asp	gga Gly	aaa Lys	gga Gly 30	aag Lys	cct Pro		96
tac ga Tyr Gl	g ggg u Gly 35	gag Glu	cag Gln	acg Thr	gta Val	aag Lys 40	ctc Leu	act Thr	gtc Val	acc Thr	aag Lys 45	ggc Gly	gga Gly	cct Pro	1	144
ctg cc Leu Pro	a ttt o Phe	gct Ala	tgg Trp	gat Asp	att Ile 55	cta Leu	tca Ser	cca Pro	cag Gln	agt Ser 60	cag Gln	tac Tyr	gga Gly	agc Ser		192
ata cca Ile Pro 65	a ttc o Phe	acc Thr	aag Lys	tac Tyr 70	cct Pro	gaa Glu	gac Asp	atc Ile	cct Pro 75	gac Asp	tat Tyr	gta Val	aag Lys	cag Gln 80	2	240
tca tto Ser Pho	c cct e Pro	gag Glu	gga Gly 85	tat Tyr	aca Thr	tgg Trp	gag Glu	agg Arg 90	atc Ile	atg Met	aac Asn	ttc Phe	gaa Glu 95	gat Asp	2	88
ggt gca Gly Ala	a gtg a Val	tgt Cys 100	act Thr	gtc Val	agc Ser	aat Asn	gat Asp 105	tcc Ser	agc Ser	atc Ile	caa Gln	ggt Gly 110	agc Ser	tgt Cys	3	336
ttc ato	tac Tyr 115	aat Asn	gtc Val	aag Lys	ttc Phe	tct Ser 120	ggt Gly	ttg Leu	aac Asn	ttt Phe	cct Pro 125	ccc Pro	aat Asn	gga Gly	3	884
cct gtt Pro Vai 130	L Met	caa Gln	aag Lys	aag Lys	aca Thr 135	cag Gln	ggc Gly	tgg Trp	gaa Glu	ccc Pro 140	aac Asn	act Thr	gag Glu	cgt Arg	4	132
ctc tti Leu Phe 145	gca Ala	cga Arg	gat Asp	gga Gly 150	atg Met	ctg Leu	ata Ile	gga Gly	aac Asn 155	aac Asn	ttt Phe	atg Met	gct Ala	ctg Leu 160	4	80
aag tto Lys Le	g gaa ı Glu	gga Gly	ggt Gly 165	ggt Gly	cat His	tat Tyr	ttg Leu	tgt Cys 170	gaa Glu	ttc Phe	aaa Lys	tct Ser	act Thr 175	tac Tyr	5	28
aag gca Lys Ala	a aag a Lys	aag Lys 180	cct Pro	gtg Val	atg Met	atg Met	cca Pro 185	ggg Gly	tat Tyr	cac His	tat Tyr	gtt Val 190	gac Asp	cgc Arg	5	76
aaa tto Lys Lei	gat Asp 195	gta Val	acc Thr	aat Asn	cac His	aac Asn 200	aag Lys	gat Asp	tac Tyr	act Thr	tcc Ser 205	gtt Val	gag Glu	cag Gln	6	24
tgt gga Cys Gly 210	, Ile	tcc Ser	att Ile	gca Ala	cgc Arg 215	aaa Lys	cct Pro	gtg Val	gtc Val	gcc Ala 220					6	60

<210> 76

<211> 220

<212> PRT

<213> Platygyra sp.

<400> 76

Ser Val Ile Ala Lys Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr 1  $\phantom{\bigg|}5\phantom{\bigg|}$  10  $\phantom{\bigg|}15\phantom{\bigg|}$ 

Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys Pro 20 25 30

Tyr Glu Gly Glu Gln Thr Val Lys Leu Thr Val Thr Lys Gly Gly Pro 35 40 45

Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Ser Gln Tyr Gly Ser 50 55 60

Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys Gln 65 70 75 80

Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ile Met Asn Phe Glu Asp 85 90 95

Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Ser Cys 100 105 110

Phe Ile Tyr Asn Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly 115 120 125

Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu Arg 130 135 140

Leu Phe Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu 145  $\phantom{\bigg|}150\phantom{\bigg|}155\phantom{\bigg|}160\phantom{\bigg|}$ 

Lys Leu Glu Gly Gly His Tyr Leu Cys Glu Phe Lys Ser Thr Tyr 165 170 175

Lys Ala Lys Lys Pro Val Met Met Pro Gly Tyr His Tyr Val Asp Arg 180 185 190

### 79/234

Lys	Leu	Asp	Val	Thr	Asn	His	Asn	Lys	Asp	Tyr	Thr	Ser	Val	Glu	Gln
		195					200					205			

Cys Gly Ile Ser Ile Ala Arg Lys Pro Val Val Ala 210 215 220

<210> 77

<211> 660

<212> DNA

<213> Pavona decussaca

<220>

<221> CDS

<222> (1)..(660)

< 40	)> '	77							
					tac Tyr				48
					gaa Glu 25				96
					ctc Leu				144
					tca Ser				192
					gac Asp				240
					gag Glu				288
					gat Asp 105				336
					ggt Gly				384

WO 02/070703	PCT/GB02/00928
--------------	----------------

					,										
W	O 02/0	70703	3					80/	/234						PCT/GB02/00
gtt Val 130	atg Met	caa Gln	aag Lys	aag Lys	aca Thr 135	cag Gln	ggc	tgg Trp	gta Val	ccc Pro 140	aac Asn	act Thr	gag Glu	cgt Arg	432
ı Phe															480
															528
															·576
															624
gaa Glu 210	att Ile	tcc Ser	att Ile	gca Ala	cgc Arg 215	aaa Lys	cct Pro	gtg Val	gtc Val	gcc Ala 220					660
10>	78														
11>	220														
L <b>2</b> > 1	PRT														
L3> :	Pavor	na de	ecuss	saca											
)0>	78														
Val	Ile	Ala	Lys 5	Gln	Met	Thr	Tyr	Lys 10	Val	Asn	Met	Ser	Gly 15	Thr	
. Asn	Gly	His 20	Туг	Phe	Glu	Val	Glu 25	Gly	Asp	Gly	Lys	Gly 30	Lys	Pro	
Glu	G1y 35	Glu	Gln	Thr	Val	Lys 40	Leu	Thr	Val	Thr	Glu 45	Gly	Gly	Pro	
Pro 50	Phe	Ala	Trp	Asp	Ile 55	Leu	Ser	Pro	Gln	Ser 60	Gln	Tyr	Gly	Ser	
	t gtto Val 130 c tttu Phe 5 g ttg s Leu g gca s Ala a ttg 210 c 10> 11> 12> 13> c T Glu u Pro	t gtt atg o Val Met 130 c ttt gca u Phe Ala 5 g ttg gaa s Leu Glu g gca aag s Ala Lys a ttg gat s Leu Asp 195 t gaa att s Glu Ile 210  10> 78 11> 220 12> PRT 13> Pavor 00> 78 r Val Ile 1 Asn Gly r Glu Gly 35 u Pro Phe	t gtt atg caa o Val Met Gln 130  c ttt gca cga u Phe Ala Arg 5  g ttg gaa gga s Leu Glu Gly  g gca aag aag s Ala Lys Lys 180  a ttg gat gta Leu Asp Val 195  t gaa att tcc s Glu Ile Ser 210  10> 78  11> 220  12> PRT  13> Pavona de  00> 78  r Val Ile Ala  1 Asn Gly His 20  r Glu Gly Glu 35  u Pro Phe Ala	o Val Met Gln Lys 130  c ttt gca cga gat u Phe Ala Arg Asp 5  g ttg gaa gga ggt s Leu Glu Gly Gly 165  g gca aag aag cct s Ala Lys Lys Pro 180  a ttg gat gta acc s Leu Asp Val Thr 195  t gaa att tcc att s Glu Ile Ser Ile 210  10> 78  11> 220  12> PRT  13> Pavona decuss  00> 78  r Val Ile Ala Lys 5  l Asn Gly His Tyr 20  r Glu Gly Glu Gln 35  u Pro Phe Ala Trp	t gtt atg caa aag aag o Val Met Gln Lys Lys 130  c ttt gca cga gat gga ger Phe Ala Arg Asp Gly 150  g ttg gaa gga ggt ggt se Leu Glu Gly Gly Gly 165  g gca aag aag cct gtg Ala Lys Lys Pro Val 180  a ttg gat gta acc aat Leu Asp Val Thr Asn 195  t gaa att tcc att gca se Glu Ile Ser Ile Ala 210  10> 78  11> 220  10> 78  11> 220  12> PRT  13> Pavona decussaca  00> 78  r Val Ile Ala Lys Gln 5  1 Asn Gly His Tyr Phe 20  r Glu Gly Glu Gln Thr 35  u Pro Phe Ala Trp Asp	t gtt atg caa aag aag aca o Val Met Gln Lys Lys Thr 130	t gtt atg caa aag aag aca cag oval Met Gln Lys Lys Thr Gln 135  c ttt gca cga gat gga atg ctg Phe Ala Arg Asp Gly Met Leu 150  g ttg gaa gga ggt ggt cat tat st Leu Glu Gly Gly Gly His Tyr 165  g gca aag aag cct gtg atg atg Asp Pro Val Met Met 180  a ttg gat gta acc aat cac aac s Leu Asp Val Thr Asn His Asn 195  t gaa att tcc att gca cgc aaa S Glu Ile Ser Ile Ala Arg Lys 210  10> 78  11> 220  12> PRT  13> Pavona decussaca  00> 78  r Val Ile Ala Lys Gln Met Thr 5  1 Asn Gly His Tyr Phe Glu Val 20  r Glu Gly Glu Gln Thr Val Lys 35  u Pro Phe Ala Trp Asp Ile Leu	t gtt atg caa aag aag aca cag ggc Val Met Gln Lys Lys Thr Gln Gly 135  c ttt gca cga gat gga atg ctg ata 150  g ttg gaa gga ggt ggt cat tat ttg 165  g gca aag aag cct gtg atg atg cca 185  g gca aag aag cct gtg atg atg cca 185  a ttg gat gta acc aat cac aac aag aag ct 185  a ttg gat gta acc aat cac aac aag aag ct 185  a ttg gat gta acc aat cac aac aag aag ct 185  a ttg gat gta acc aat cac aac aag aag ct 195  a ttg gaa att tcc att gca cgc aaa cct 185  a ttg gaa att tcc att gca cgc aaa cct 185  a ttg Ser Ile Ala Arg Arg Lys Pro 215  10> 78  11> 220  12> PRT  13> Pavona decussaca  00> 78  r Val Ile Ala Lys Gln Met Thr Tyr 5  1 Asn Gly His Tyr Phe Glu Val Glu 25  r Glu Gly Glu Gln Thr Val Lys Leu 35  u Pro Phe Ala Trp Asp Ile Leu Ser	t gtt atg caa aag aag aca cag ggc tgg ttt gca gag ggt ggt ggt he Glu Glu Gly 150  g ttg gaa gga ggt ggt ggt ttg the Glu Glu Gly 165  g gca aag aag cct gtg atg try Lys 170  g gca aag aag cct gtg atg atg cca ggg s Ala Lys Lys Pro Val Met Met Pro Glu 185  a ttg gat gta acc aat cac aac aag gat S Leu Asp Val Thr Asn His Asn Lys Asp 195  t gaa att tcc att gca cgc aaa cct gtg S Glu Ile Ser Ile Ala Arg Lys Pro Val 215  10> 78  11> 220  12> PRT  13> Pavona decussaca  00> 78  r Val Ile Ala Lys Gln Met Thr Tyr Lys 5  1 Asn Gly His Tyr Phe Glu Val Glu Gly 25  r Glu Gly Glu Gln Thr Val Lys Leu Thr 35  u Pro Phe Ala Trp Asp Ile Leu Ser Pro	## 130   Section   Section	## 180/234  ## 18	t gtt gtt atg Caa aag aag gga ggt ggt cat ttg gaa aag aag aag atg Caa aag gat lac Caa aag gac aag aag aag aag aag aag aag a	t gtt atg Caa aag aag gga ggg gga ggt ggg atg gga aac tct atg She Ala Lys Lys Thr 130	t gtt atg caa aag aag atg gaa atg ctg atg gaa atc aac tta gtg gaa atg s Leu Spy Spr Val Met Slu Spy	## 180/234  It git atg caa aag aag aca cag ggc tgg gta ccc aac act gag cgt can lead lead by a list of land lead of land lead to val met Gln Lys Lys Thr 135  It git gaa gga ggt ggd cat tat tig gga aac acc tig acc acc act gag cgt gg tag pro he Ala Arg Asp Gly Met 150  It gaa gga ggt ggt cat tat tig gtg gaa tig acc acc act tacc acc acc gg gas gg ggt ggt cat tat tig gtg gaa tig acc acc acc acc acc acc acc acc acc gg gas gg ggt ggt cat tat tig gtg gaa tig acc acc acc acc acc acc acc acc acc ac

Val Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys Gln 65 70 75 80

Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ile Met Asn Phe Glu Asp  $85 \\ 90 \\ 95$ 

#### 81/234

Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys 100 105 110

Phe Ile Tyr Asn Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly

Phe He Tyr Asn Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly
115 120 125

Pro Val Met Gln Lys Lys Thr Gln Gly Trp Val Pro Asn Thr Glu Arg 130 135 140

Leu Phe Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu 145 150 155 160

Lys Leu Glu Gly Gly His Tyr Leu Cys Glu Phe Lys Ser Thr Tyr 165 170 175

Lys Ala Lys Lys Pro Val Met Met Pro Gly Tyr His Tyr Val Asp Arg 180 185 190

Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu Gln
195 200 205

Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 210 215 220

<210> 79

<211> 660

<212> DNA

<213> Pavona decussaca

<220>

<221> CDS

<222> (1)..(660)

<400> 79

tcc gtt atc gct aaa cag atg acc tac aag gtt tat atg tca ggc acg Ser Val Ile Ala Lys Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr 1 5 10 15

gtc aat gga cac tac ttt gag gtc gaa ggc gat gga aaa gga aag cct 96 Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys Pro 20 25 30

48

WO 02/070703		PCT/GB02/00928
	82/234	

tac Tyr	gag Glu	ggg Gly 35	gag Glu	cag Gln	acg Thr	gta Val	aag Lys 40	ctc Leu	act Thr	gtc Val	acc Thr	aag Lys 45	ggc Gly	gga Gly	cct Pro	144
ctg Leu	cca Pro 50	ttt Phe	gct Ala	tgg Trp	gat Asp	att Ile 55	tta Leu	tca Ser	cca Pro	cag Gln	tgt Cys 60	cag Gln	tac Tyr	gga Gly	agc Ser	192
ata Ile 65	cca Pro	ttc Phe	acc Thr	aag Lys	tac Tyr 70	cct Pro	gaa Glu	gac Asp	atc Ile	cct Pro 75	gac Asp	tat Tyr	gta Val	aag Lys	cag Gln 80	240
tca Ser	ttc Phe	ccg Pro	gag Glu	gga Gly 85	ttt Phe	aca Thr	tgg Trp	gag Glu	agg Arg 90	atc Ile	atg Met	aac Asn	ttt Phe	gaa Glu 95	gat Asp	288
ggt Gly	gca Ala	gtg Val	tgt Cys 100	act Thr	gtc Val	agc Ser	aat Asn	gat Asp 105	tcc Ser	agc Ser	atc Ile	caa Gln	ggc Gly 110	aac Asn	tgt Cys	336
ttc Phe	acc Thr	tac Tyr 115	cat His	gtc Val	aag Lys	ttc Phe	tct Ser 120	ggt Gly	ttg Leu	aac Asn	ttt Phe	cct Pro 125	ccc Pro	aat Asn	GJÀ āāā	384
cct Pro	gtg Val 130	atg Met	cag Gln	aag Lys	aag Lys	aca Thr 135	cag Gln	ggc Gly	tgg Trp	gaa Glu	ccc Pro 140	cac His	tct Ser	gag Glu	cgt Arg	432
ctc Leu 145	ttt Phe	gca Ala	cgg Arg	ggt Gly	gga Gly 150	atg Met	ctg Leu	ata Ile	gga Gly	aac Asn 155	aac Asn	ttt Phe	atg Met	gct Ala	ccg Pro 160	480
aag Lys	tta Leu	gaa Glu	gga Gly	ggc Gly 165	ggt Gly	cac His	tat Tyr	ttg Leu	tgt Cys 170	gaa Glu	ttc Phe	aaa Lys	act Thr	act Thr 175	tac Tyr	528
aag Lys	gca Ala	aag Lys	aag Lys 180	cct Pro	gtg Val	aag Lys	atg Met	ccg Pro 185	ggg Gly	tat Tyr	cat His	tat Tyr	gtt Val 190	gac Asp	cgc Arg	576
aaa Lys	ctg Leu	gat Asp 195	gta Val	acc Thr	aat Asn	cac His	aac Asn 200	aag Lys	gat Asp	tac Tyr	act Thr	tcc Ser 205	gtt Val	gag Glu	cag Gln	624
tgt Cys	gaa Glu 210	atc Ile	tcc Ser	att Ile	gca Ala	cgc Arg 215	aaa Lys	cct Pro	gtg Val	gtc Val	gcc Ala 220					660

<210> 80

<211> 220

<212> PRT

<213> Pavona decussaca

<400> 80

#### WO 02/070703 PCT/GB02/00928 83/234

---

Ser Val Ile Ala Lys Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr 1 5 10 15

Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys Pro 20 25 30

Tyr Glu Gly Glu Gln Thr Val Lys Leu Thr Val Thr Lys Gly Gly Pro 35 40 45

Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Cys Gln Tyr Gly Ser 50 60

Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys Gln 65 70 75 80

Ser Phe Pro Glu Gly Phe Thr Trp Glu Arg Ile Met Asn Phe Glu Asp 85 90 95

Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys 100 105 110

Phe Thr Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly 115 120 125

Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro His Ser Glu Arg 130 135 140

Leu Phe Ala Arg Gly Gly Met Leu Ile Gly Asn Asn Phe Met Ala Pro 145 150 155

Lys Leu Glu Gly Gly His Tyr Leu Cys Glu Phe Lys Thr Thr Tyr 165 170 175

Lys Ala Lys Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Asp Arg 180 185 190

Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu Gln 195 200 205

Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 210 215 220

<210> 81

<211> 660

84/234

<212> DNA <213> Pavona decussaca <220> <221> CDS <222> (1)..(660) <400> 81 tee gtt ate get aaa eag atg ace tae aag gtt tat atg tea gge acg 48 Ser Val Ile Ala Lys Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr gtc aat gga cac tac ttt gag gtt gaa ggc gat gga aaa gga aaq cct 96 Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys Pro tac gag ggg gag cag acg gta aag ctc act gtc acc gag ggc gga cct 144 Tyr Glu Gly Glu Gln Thr Val Lys Leu Thr Val Thr Glu Gly Gly Pro 35 40 ctg cca ttt gct tgg gat att cta tca cca cag agt cag tac gga agc 192 Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Ser Gln Tyr Gly Ser 50 gta cca ttc acc aag tac cct gaa gac atc cct gac tat gta aag cag 240 Val Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys Gln tca ttc ccg gag gga tat aca tgg gag agg atc atg aac ttt gaa gat 288 Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ile Met Asn Phe Glu Asp ggt gca gtg tgt act gtc agc aat gat tcc agc atc caa ggc aac tgt 336 Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys ttc atc tac cat gtc aag ttc tct ggt ttg aac ttt cct ccc aat gga 384 Phe Ile Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly 120 cct gtt atg cag aag aca cag ggc tgg gaa ccc aac act gag cgt 432 Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu Arg 135 ctc tct gca cga gat gga atg cta ata gga aac aac ttt atg gct ctg 480 Leu Ser Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu 150 155 aag tta gaa gga ggt ggt cac tat ttg tgt gaa ttc aaa tct act tac 528 Lys Leu Glu Gly Gly His Tyr Leu Cys Glu Phe Lys Ser Thr Tyr

170

#### 85/234

									85/2	234						
aag Lys	gca Ala	agg Arg	aag Lys 180	cct Pro	gtg Val	aag Lys	atg Met	cca Pro 185	Gly ggg	tat Tyr	cac His	tat Tyr	gtt Val 190	gac Asp	cgc Arg	576
aaa Lys	ctg Leu	gat Asp 195	gta Val	acc Thr	aat Asn	cac His	aac Asn 200	aag Lys	gat Asp	tac Tyr	act Thr	tcc Ser 205	gtt Val	gag Glu	cag Gln	624
cgt Arg	gaa Glu 210	att Ile	tcc Ser	att Ile	gca Ala	cgc Arg 215	aaa Lys	cct Pro	gtg Val	gtc Val	gcc Ala 220					660
<210	)> :	82														
<211	l> :	220						٠								
<212	2>	PRT														
<213	3> :	Pavor	na de	cus	saca											
<400	)> :	32														
Ser 1	Val	Ile	Ala	Lys 5	Gln	Met	Thr	Tyr	Lys 10	Val	Tyr	Met	Ser	Gly 15	Thr	
Val	Asn	Gly	His 20	Tyr	Phe	Glu	Val	Glu 25	Gly	Asp	Gly	Lys	Gly 30	Lys	Pro	
Tyr	Glu	Gly 35	Glu	Gln	Thr	Val	Lys 40	Leu	Thr	Val	Thr	Glu 45	Gly	Gly	Pro	
Leu	Pro 50	Phe	Ala	Trp	Asp	Ile 55	Leu ·	Ser	Pro	Gln	Ser 60	Gln	Tyr	Gly	Ser	
Val 65	Pro	Phe	Thr	Lys	Tyr 70	Pro	Glu	Asp	Ile	Pro 75	Asp	Tyr	Val	Lys	Gln 80	
Ser	Phe	Pro	Glu	Gly 85	Tyr	Thr	Trp	Glu	Arg 90	Ile	Met	Asn	Phe	Glu 95	Asp	
Gly	Ala	Val	Cys	Thr	Val	Ser	Asn	Asp	Ser	Ser	Ile	Gln	Gly	Asn	Cys	

100 105 110

Phe Ile Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly

Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu Arg

120

115

	r Ala	Arg	Asp	Gly 150	Met	Leu	Ile	Gly	Asn 155	Asn	Phe	Met	Ala	Leu 160	
Lys Le	u Glu	Gly	Gly 165	Gly	His	Tyr	Leu	Cys 170	Glu	Phe	Lys	Ser	Thr 175	Tyr	
Lys Al	a Arg	Lys 180	Pro	Val	Lys	Met	Pro 185	Gly	Tyr	His	Tyr	Val 190	Asp	Arg	
Lys Le	195	Val	Thr	Asn	His	Asn 200	Lys	Asp	Tyr	Thr	Ser 205	Val	Glu	Gln	
Arg Gl 21		Ser	Ile	Ala	Arg 215	Lys	Pro	Val	Val	Ala 220					
<210>	83														
<211>	600														
<212>	DNA														
<213>	Mont:	ipora	asp.												
<220>															
<221>															
12217	CDS														
<222>	(1).	. (600	))												
<222> <400>	(1). 83														
<222>	(1). 83	gct	aaa	cag Gln	atg Met	acc Thr	tac Tyr	aag Lys 10	gtt Val	tat Tyr	atg Met	tca Ser	ggc Gly 15	acg Thr	48
<222> <400> tcc gt	(1). 83 atc Ile	gct Ala	aaa Lys 5	Gln act	Met gtc	Thr	Tyr	Lys 10 ggc	Val gga	Tyr cct	Met	Ser	Gly 15 ttt	Thr	48 96
<222> <400> tcc gt: Ser Va: 1 gtc aai	83 atc Ile gga Gly	gct Ala cac His 20	aaa Lys 5 ctc Leu	Gln act Thr	Met gtc Val cag	acc Thr	Tyr aag Lys 25 cag	Lys 10 ggc Gly tac	Val gga Gly	Tyr cct Pro	Met ctg Leu ata	cca Pro 30	Gly 15 ttt Phe ttc	Thr gct Ala acc	
<222> <400> tcc gt: Ser Va: 1 gtc aai Val Asi	83 atc Ile gga Gly att Ile 35 cct	gct Ala cac His 20 cta Leu	aaa Lys 5 ctc Leu tca Ser	Gln act Thr cca Pro	Met gtc Val cag Gln cct	Thr acc Thr agt Ser 40	Tyr aag Lys 25 cag Gln	Lys 10 ggc Gly tac Tyr	yal gga Gly gga Gly	Tyr cct Pro agc Ser	Met ctg Leu ata Ile 45 tca	cca Pro 30 cca Pro	Gly 15 ttt Phe ttc Phe	Thr gct Ala acc Thr	96

						,										
	W	02/0	70703	3					97/	234						PCT/GB02/009
act Thr	gtc Val	agc Ser	aat Asn	gat Asp 85	tcc Ser	agc Ser	atc Ile	caa Gln	ggt Gly 90	aac Asn	tgt Cys	ttc Phe	atc Ile	tac Tyr 95	aat Asn	288
gtc Val	aag Lys	ttc Phe	tct Ser 100	ggt Gly	ttg Leu	aac Asn	ttt Phe	cct Pro 105	ccc Pro	aat Asn	gga Gly	cct Pro	gtt Val 110	atg Met	caa Gln	336
aag Lys	aag Lys	aca Thr 115	cag Gln	ggc Gly	tgg Trp	gaa Glu	ccc Pro 120	aac Asn	act Thr	gag Glu	cgt Arg	ctc Leu 125	ttt Phe	gca Ala	cga Arg	384
gat Asp	gga Gly 130	atg Met	ctg Leu	ata Ile	gga Gly	aac Asn 135	aac Asn	ttt Phe	atg Met	gct Ala	ctg Leu 140	aag Lys	ttg Leu	gaa Glu	gga Gly	432
ggt Gly 145	ggt Gly	cat His	tat Tyr	ttg Leu	tgt Cys 150	gaa Glu	ttc Phe	aaa Lys	tct Ser	act Thr 155	tac Tyr	aag Lys	gca Ala	aag Lys	aag Lys 160	480
cct Pro	gtg Val	atg Met	atg Met	cca Pro 165	ggg Gly	tat Tyr	cac His	tat Tyr	gtt Val 170	gac Asp	cgc Arg	aaa Lys	ttg Leu	gat Asp 175	gta Val	528
acc Thr	aat Asn	cac His	aac Asn 180	aag Lys	gat Asp	tac Tyr	act Thr	tcc Ser 185	gtt Val	gag Glu	cag Gln	tgt Cys	gaa Glu 190	att Ile	ccc Pro	576
		cgc Arg 195			_											600
<21	)> {	34														
<21	1> 2	200														
<212	2> I	PRT														
<21	3> 1	4onti	.pora	sp.												
<400	)> 8	34														
Ser 1	Val	Ile	Ala	Lys 5	Gln	Met	Thr	Туг	Lys 10	Val	Tyr	Met	Ser	Gly 15	Thr	

Val Asn Gly His Leu Thr Val Thr Lys Gly Gly Pro Leu Pro Phe Ala 20 25 30

Trp Asp Ile Leu Ser Pro Gln Ser Gln Tyr Gly Ser Ile Pro Phe Thr 35 40 45

Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys Gln Ser Phe Pro Glu 50 60

#### 88/234

Gly Tyr Thr Trp Glu Arg Ile Met Asn Phe Glu Asp Gly Ala Val Cys
65 70 75

Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys Phe Ile Tyr Asn 85 90 95

Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly Pro Val Met Gln
100 105 110

Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu Arg Leu Phe Ala Arg 115 120 125

Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu Lys Leu Glu Gly 130 135

Gly Gly His Tyr Leu Cys Glu Phe Lys Ser Thr Tyr Lys Ala Lys Lys 145 150 155

Pro Val Met Met Pro Gly Tyr His Tyr Val Asp Arg Lys Leu Asp Val 165 170 175

Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu Gln Cys Glu Ile Pro 180 185 190

Ile Ala Arg Lys Pro Val Val Ala 195 200

<210> 85

<211> 660

<212> DNA

<213> Montipora sp.

<220>

<221> CDS

<222> (1)..(660)

<400> 85

tcc gtt atc gct aaa cag atg acc tac aag gtt tat atg tca ggc acg Ser Val Ile Ala Lys Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr 1 10 15

48

89/234

			cac His 20											96
			gag Glu											144
			gct Ala											192
			acc Thr											240
			gag Glu											288
			tgt Cys 100											336
			aat Asn											384
	_	_	caa Gln	•	_		_	 	_				_	432
		-	cga Arg	_		_	_				_	_	_	480
			gga Gly											528
_	-		aag Lys 180			_	_				_	_	-	576
			gta Val											624
			tcc Ser											660

<210> 86

<211> 220

<212> PRT

<213> Montipora sp.

<400> 86

Ser Val Ile Ala Lys Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr

Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys Pro

Tyr Glu Gly Glu Gln Thr Val Lys Leu Thr Val Thr Lys Gly Gly Pro

Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Ser Gln Tyr Gly Ser 55

Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys Gln 70

Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ile Met Asn Phe Glu Asp 85 90 -

Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys 100

Phe Ile Tyr Asn Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly

Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu Arg 135 130

Leu Phe Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu

Lys Leu Glu Gly Gly His Tyr Leu Cys Glu Phe Lys Ser Thr Tyr 165 170

Lys Ala Lys Lys Pro Val Met Met Pro Gly Tyr His Tyr Val Asp Arg 180 185

Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu Gln 200

Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 215

<210>	87														
<211>	660														
<212>	DNA														
<213>	Mont.	ipora	a sp	•											
	•														
<220>															
<221>															
<222>															
<400> tcc gt	87 tatc	act	aaa	caq	atq	acc	tac	aag	att	tat	atq	tca	aac	aca	48
Ser Va															
gtc aa Val As															96
		20					25	1		~_ <b>J</b>	-,,-	30	_, _		
tac ga Tyr Gl			_	-	-	-			-		_				144
	35					40					45				
ctg cc	o Phe	_		-	Ile				_	Ser	_				192
50					55					60					
ata co			-	Tyr		_	_		Pro	-		-	-	Gln	240
65	c cct	<b>~</b> ~~	~~~	70	2.02	+~~	<b>~~~</b>	200	75 2± a	2+4		++-	~~~	80	200
tca tt Ser Ph															288
ggt gc	a ata	tat		atc	agc	aat	gat		agc	atc	caa	aat		tat	336
Gly Al															
ttc at															384
Phe Il	e Tyr 115	Asn	Val	Lys	Phe	Ser 120	Gly	Leu	Asn	Phe	Pro 125	Pro	Asn	Gly	
cct gt															432
Pro Va		GIN	гла	гуз	135	GIU	стÀ	rtb	GIU	140	ASN	ınr	GIU	Arg	
ctc tt Leu Ph															480
145		9	F	150				2	155					160	

#### 92/234

_	_	_		-	ggt Gly		_	-	_					528
_	_	_	_		gtg Val	_					-	_	-	576
	_	_	-		aat Asn		_	_			_		_	624
_	_				gca Ala	-		_	_	_			۰.	660

<210> 88

<211> 220

<212> PRT

<213> Montipora sp.

<400> 88

Ser Val Ile Ala Lys Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr 1  $\phantom{\bigg|}$  5  $\phantom{\bigg|}$  10  $\phantom{\bigg|}$  15

Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys Pro 20 25 30

Tyr Glu Gly Glu Gln Thr Val Lys Leu Thr Val Thr Lys Gly Gly Pro 35 40 45

Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Ser Gln Tyr Gly Ser 50 55 60

Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys Gln 65 70 75 80

Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ile Met Asn Phe Glu Asp 85 . 90 95

Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys 100 105 110

Phe Ile Tyr Asn Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly 115 120 125

130	n Lys Lys	Thr Gln 135	Gly Trp	Glu Pro 140	Asn Thr	Glu Arg	
Leu Phe Ala Ar 145	g Asp Gly 150	Met Leu	Ile Gly	Asn Asn 155	Phe Met	Ala Leu 160	
Lys Leu Glu Gl	y Gly Gly 165	His Tyr	Leu Cys 170	Glu Phe	Lys Ser	Thr Tyr 175	
Lys Ala Lys Ly 18		Met Met	Pro Gly 185	Tyr His	Tyr Val 190	Asp Arg	
Lys Leu Asp Va 195	l Thr Asn	His Asn 200	Lys Asp	Tyr Thr	Ser Val 205	Glu Gln	
Cys Glu Ile Se 210	r Ile Ala	Arg Lys 215	Pro Val	Val Ala 220			
<210> 89							
<211> 765							
<212> DNA							
40725 B							
<213> Acropor	a aspera		•		•		
<213> Acropor	a aspera						
	a aspera						
<220>						•	
<220> <221> CDS						•	
<220> <221> CDS <222> (1)(7	- 65) t ttg cgt	gta atg Val Met	gac atc Asp Ile 10	agc atc Ser Ile	tct ttc Ser Phe	acg gaa Thr Glu 15	48
<220> <221> CDS <222> (1)(7) <400> 89 gcg acc aca gg Ala Thr Thr Gl	65) t ttg cgt y Leu Arg 5 a gaa acg	Val Met	Asp Ile 10 aat cgt	Ser Ile	Ser Phe	Thr Glu 15 ctt att	48
<220> <221> CDS <222> (1)(7) <400> 89 gcg acc aca gg Ala Thr Thr Gl 1 gga gct act ta	t ttg cgt y Leu Arg 5 a gaa acg Glu Thr 20 t gtg atc	Val Met ttt gcg Phe Ala gct aca	Asp Ile 10 aat cgt Asn Arg 25 caa atg	Ser Ile tgt tct Cys Ser acc tac	Ser Phe gcg cta Ala Leu aag gtt	Thr Glu 15 ctt att Leu Ile 30 tat atg 1	

WO 02/070703	PCT/GB02/00928
--------------	----------------

94/234

gga Gly	aag Lys 65	cct Pro	tac Tyr	gag Glu	G] À aaa	gag Glu 70	cag Gln	acg Thr	gta Val	agg Arg	ctg Leu 75	gct Ala	gtc Val	acc Thr	aag Lys	240
ggc Gly	gga Gly	cct Pro	ctg Leu	cca Pro	ttt Phe 85	gct Ala	tgg Trp	gat Asp	att Ile	tta Leu 90	tca Ser	cca Pro	cag Gln	tgt Cys	cag Gln 95	288
			ata Ile													336
gta Val	aag Lys	cag Gln	tca Ser 115	ttc Phe	ccg Pro	gag Glu	gga Gly	ttt Phe 120	aca Thr	tgg Trp	gag Glu	agg Arg	atc Ile 125	atg Met	aac Asn	384
ttt Phe	gaa Glu	gat Asp 130	ggt Gly	gca Ala	gtg Val	tgt Cys	act Thr 135	gtc Val	agc Ser	aat Asn	gat Asp	tcc Ser 140	agc Ser	atc Ile	caa Gln	432
ggc Gly	aac Asn 145	tgt Cys	ttc Phe	atc Ile	tac Tyr	cat His 150	gtc Val	aag Lys	ttc Phe	tct Ser	ggt Gly 155	ttg Leu	aac Asn	ttt Phe	cct Pro	480
ccc Pro 160	aat Asn	gga Gly	cct Pro	gtt Val	atg Met 165	cag Gln	aag Lys	aag Lys	aca Thr	cag Gln 170	ggc Gly	tgg Trp	gaa Glu	ccc Pro	cac His 175	528
			ctc Leu													576
atg Met	gct Ala	ctg Leu	aag Lys 195	tta Leu	gaa Glu	gga Gly	ggc Gly	ggt Gly 200	cac His	tat Tyr	ttg Leu	tgt Cys	gaa Glu 205	ttc Phe	aaa Lys	624
act Thr	act Thr	tac Tyr 210	aag Lys	gca Ala	aag Lys	aag Lys	cct Pro 215	gtg Val	aag Lys	atg Met	cca Pro	ggg Gly 220	tat Tyr	cat His	tat Tyr	672
gtt Val	gac Asp 225	cgc Arg	aaa Lys	ctg Leu	gat Asp	gta Val 230	acc Thr	aat Asn	cac	aac Asn	aag Lys 235	gat Asp	tac Tyr	act Thr	tcc Ser	720
gtt Val 240	Glu	cag Gln	tgt Cys	gaa Glu	att Ile 245	Ser	att Ile	aca Thr	cgc Arg	aaa Lys 250	Pro	gtg Val	gtc Val	gcc Ala		765

<210> 90

<211> 19

<212> PRT

<213> Acropora aspera

95/234

Ala Thr Thr Gly Leu Arg Val Met Asp Ile Ser Ile Ser Phe Thr Glu
1 5 10 15

Gly Ala Thr

<210> 91

<211> 235

<212> PRT

<213> Acropora aspera

<400> 91

Glu Thr Phe Ala Asn Arg Cys Ser Ala Leu Leu Ile Leu Asn Met Ser 1 5 10 15

Val Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr Val 20  $\cdot$  25 30

Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys Pro Tyr 35 40 45

Glu Gly Glu Gln Thr Val Arg Leu Ala Val Thr Lys Gly Gly Pro Leu 50 60

Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Cys Gln Tyr Gly Ser Ile 65. 70 75 80

Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys Gln Ser 85 90 95

Phe Pro Glu Gly Phe Thr Trp Glu Arg Ile Met Asn Phe Glu Asp Gly 100 105 110

Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys Phe 115 120 125

Ile Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly Pro 130 135 140

Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro His Ser Glu Arg Leu 145 150 155 160

		96/234	
	Gly Met Leu Ile Gly 165	Asn Asn Phe Met Al 170	a Leu Lys 175
Leu Glu Gly Gly G 180	Gly His Tyr Leu Cys 185	Glu Phe Lys Thr Th 19	
Ala Lys Lys Pro V 195	al Lys Met Pro Gly 200	Tyr His Tyr Val As 205	p Arg Lys
Leu Asp Val Thr A	Asn His Asn Lys Asp 215	Tyr Thr Ser Val Gl 220	u Gln Cys
Glu Ile Ser Ile T 225	Chr Arg Lys Pro Val 230	Val Ala 235	
<210> 92			
<211> 765			
<212> DNA			
<213> Acropora a	aspera		
<220>			
<221> CDS			
<222> (1)(765)			•
<400> 92			
gcg acc aca ggt t	eu Arg Val Met Asp	atc agc atc tct tt Ile Ser Ile Ser Ph 10	
Gly Ala Thr G	gaa acg ttt gcg aat ilu Thr Phe Ala Asn 0	cgt tgt tct gcg ct Arg Cys Ser Ala Le 25	a ctt att 96 u Leu Ile 30
ctc aat atg agt g Leu Asn Met Ser V 35	tg atc gct aca caa al Ile Ala Thr Gln 40	atg acc tac aag gt Met Thr Tyr Lys Va. 45	t tat atg 144 l Tyr Met
		gag gtc gaa ggc ga Glu Val Glu Gly As 60	
gga aag oot tac g Gly Lys Pro Tyr G 65	ag ggg gag cag acg lu Gly Glu Gln Thr 70	gta agg ctg gct gtc Val Arg Leu Ala Va 75	c acc aag 240 L Thr Lys

WO 02/070703 PCT/G	B02/00928
--------------------	-----------

97/234

ggc Gly 80	gga Gly	cct Pro	ctg Leu	cca Pro	ttt Phe 85	gct Ala	tgg Trp	gat Asp	att Ile	tta Leu 90	tca Ser	cca Pro	cag Gln	tgt Cys	cag Gln 95	288
tac Tyr	gga Gly	agc Ser	ata Ile	cca Pro 100	ttc Phe	acc Thr	aag Lys	tac Tyr	cct Pro 105	gaa Glu	gac Asp	atc Ile	cct Pro	gac Asp 110	tat Tyr	336
gta Val	aag Lys	cag Gln	tca Ser 115	ttc Phe	ccg Pro	gag Glu	gga Gly	ttt Phe 120	aca Thr	tgg Trp	gag Glu	agg Arg	atc Ile 125	atg Met	aac Asn	384
			ggt Gly													432
			ttc Phe													480
			cct Pro													528
			ctc Leu													576
			aag Lys 195													624
			aag Lys													672
			aaa Lys													720
-		_	tgt Cys	_					-				_	_		765
<b>~21</b> (	)	7.3														

<210> 93

<211> 19

<212> PRT

<213> Acropora aspera

<400> 93

Ala Thr Thr Gly Leu Arg Val Met Asp Ile Ser Ile Ser Phe Thr Glu 1 5 10 15

PCT/GB02/00928

98/234

Gly Ala Thr

<210> 94

<211> 235

<212> PRT

<213> Acropora aspera

<400> 94

Glu Thr Phe Ala Asn Arg Cys Ser Ala Leu Leu Ile Leu Asn Met Ser

Val Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr Val

Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys Pro Tyr

Glu Gly Glu Gln Thr Val Arg Leu Ala Val Thr Lys Gly Gly Pro Leu 50 55

Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Cys Gln Tyr Gly Ser Ile

Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys Gln Ser

Phe Pro Glu Gly Phe Thr Trp Glu Arg Ile Met Asn Phe Glu Asp Gly 105

Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys Phe 120

Ile Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly Pro 130

Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro His Ser Glu Arg Leu 150 155

Phe Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu Lys 170

Leu Glu	Gly	Gly 180	Gly	His	Tyr	Leu	Cys 185	Glu	Phe	Lys	Thr	Thr 190	Tyr	Lys	
Ala Lys	Lys 195	Pro	Val	Lys	Met	Pro 200	Gly	Tyr	His	Tyr	Val 205	Asp	Arg	Lys	
Leu Asp 210		Thr	Asn	His	Asn 215	Lys	Asp	Tyr	Thr	Ser 220	Val	Glu	Gln	Cys	
Glu Ile 225	Ser	Ile	Thr	Arg 230	Lys	Pro	Val	Val	Ala 235						
<210>	95														
<211>	660														
<212>	DNA														
<213>	Acro	pora	aspe	era											
<220>															
<221>	CDS														
<222>	(1).	. (660	0)												
<400>	95	ant			-+-		+		~++	+-+					40
agt ggg Ser Gly 1															48
gtc aat Val Asn															96
tac gag Tyr Glu															144
ctg cca Leu Pro 50		_		-					_	_	_			-	192
ata cca Ile Pro 65						-	-			-		_	_	•	240
tca tto Ser Phe															288

WO 02/070703	3			PCT/GB6	02/00928
		100	/234		
ggt gca gtg tgt Gly Ala Val Cys 100	Thr Val Ser	aat gat tcc Asn Asp Ser 105	agc atc caa ggc Ser Ile Gln Gly 110	aac tgt 3 Asn Cys	36
ttc atc tac cat Phe Ile Tyr His 115	gtc aag ttc Val Lys Phe	tct ggt ttg Ser Gly Leu 120	aac ttt cct ccc Asn Phe Pro Pro 125	aat gga 3 Asn Gly	84
cct gtt atg cag Pro Val Met Gln 130	aag aag aca Lys Lys Thr 135	cag ggc tgg Gln Gly Trp	gaa ccc cac tct Glu Pro His Ser 140	gag cgt 4 Glu Arg	32
ctc ttt gca cga Leu Phe Ala Arg 145	gac gga atg Asp Gly Met 150	ctg ata gga Leu Ile Gly	aac aac ttt atg Asn Asn Phe Met 155	gct ctg 4 Ala Leu 160	80
aag tta gaa gga Lys Leu Glu Gly	ggc ggt cac Gly Gly His 165	tat ttg tgt Tyr Leu Cys 170	gaa ttc aaa act Glu Phe Lys Thr	act tac 5 Thr Tyr 175	28
aag gca aag aag Lys Ala Lys Lys 180	Pro Val Lys	atg cca ggg Met Pro Gly 185	tat cat tat gtt Tyr His Tyr Val 190	gac cgc 5 Asp Arg	76
aaa ctg gat gta Lys Leu Asp Val 195	atc aat cac Ile Asn His	aac aag gat Asn Lys Asp 200	tac act tcc gtt Tyr Thr Ser Val 205	gag cag 6 Glu Gln	24
tgt gaa att tcc Cys Glu Ile Ser 210				6	60
<210> 96					
<211> 220					
<212> PRT					
<213> Acropora	aspera				
<400> 96					
Ser Gly Ile Ala	Thr Gln Met 5	Thr Tyr Lys	Val Tyr Met Ser	Gly Thr 15	
Val Asn Gly His	Tyr Phe Glu	Val Glu Gly 25	Asp Gly Lys Gly 30	Lys Pro	
Tyr Glu Gly Glu 35	Gln Thr Val	Arg Leu Ala 40	Val Thr Lys Gly 45	Gly Pro	
Leu Pro Phe Ala	Trp Asp Ile	Leu Ser Pro	Gln Cys Gln Tyr 60	Gly Ser	

#### 101/234

Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys Gln 65 70 75 80

Ser Phe Pro Glu Gly Phe Thr Trp Glu Arg Ile Met Asn Phe Glu Asp 85 90 95

Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys 100 105 110

Phe Ile Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly 115 120 125

Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro His Ser Glu Arg 130 135 140

Leu Phe Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu 145 150 155 160

Lys Leu Glu Gly Gly His Tyr Leu Cys Glu Phe Lys Thr Tyr 165 170 175

Lys Ala Lys Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Asp Arg

Lys Leu Asp Val Ile Asn His Asn Lys Asp Tyr Thr Ser Val Glu Glu 195 200 205

Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 210 215 220

<210> 97

<211> 660

<212> DNA

<213> Acropora aspera

<220>

<221> CDS

<222> (1)..(660)

<400> 97

WO 02/070703		PCT/GB02/00928
	100/024	

	***	U 112/1	,,,,,,,,	,					102	2/234						1 € 17 € 10.27 000
agt Ser 1	gtg Val	atc Ile	gct Ala	aca Thr 5	caa Gln	atg Met	acc Thr	tac Tyr	aag Lys 10	gtt Val	tat Tyr	atg Met	tca Ser	ggc Gly 15	acg Thr	48
gtc Val	aat Asn	gga Gly	cac His 20	tac Tyr	ttt Phe	gag Glu	gtc Val	gaa Glu 25	ggc Gly	gat Asp	gga Gly	aaa Lys	gga Gly 30	aag Lys	cct Pro	96
tac Tyr	gag Glu	ggg Gly 35	gag Glu	cag Gln	acg Thr	gta Val	agg Arg 40	ctg Leu	gct Ala	gtc Val	acc Thr	aag Lys 45	ggc Gly	gga Gly	cct Pro	144
ctg Leu	cca Pro 50	ttt Phe	gcc Ala	tgg Trp	gat Asp	att Ile 55	tta Leu	tca Ser	cca Pro	cag Gln	tgt Cys 60	cag Gln	tac Tyr	gga Gly	agc Ser	192
ata Ile 65	cca Pro	ttc Phe	acc Thr	aag Lys	tac Tyr 70	cct Pro	gaa Glu	gac Asp	atc Ile	cct Pro 75	gac Asp	tat Tyr	gta Val	aag Lys	cag Gln 80	240
	ttc Phe															288
	gca Ala															336
	atc Ile			_	_				_							384
	gtt Val 130															432
	ttt Phe															480
	tta Leu															528
_	gca Ala	_	_			_	_						-		_	576
	ctg Leu															624
tgt Cys	gaa Glu 210	att Ile	tcc Ser	att Ile	gca Ala	cgc Arg 215	aac Asn	cct Pro	gtg Val	gtc Val	gcc Ala 220					660

<210> 98

<211> 220

<212> PRT

<213> Acropora aspera

<400> 98

Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys Pro 20 25 30

Tyr Glu Gly Glu Gln Thr Val Arg Leu Ala Val Thr Lys Gly Gly Pro 35 40 45

Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Cys Gln Tyr Gly Ser 50 60

Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys Gln 65 70 75 80

Ser Phe Pro Glu Gly Phe Thr Trp Glu Arg Ile Met Asn Phe Glu Asp 85 90 95

Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys 100 105 110

Phe Ile Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly 115 120 125

Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro His Ser Glu Arg 130 135 140

Leu Phe Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu 145 150 155

Lys Leu Glu Gly Gly Gly His Tyr Leu Cys Glu Phe Lys Thr Tyr 165 170 175

Lys Ala Lys Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Asp Arg 180 185 190

Lys Leu Asp Val Ile Asn His Asn Lys Asp Tyr Thr Ser Val Glu Gln 195 200 205

Cys	Glu 210	Ile	Ser	Ile	Ala	Arg 215	Asn	Pro	·Val	Val	Ala 220					
<210	)> !	99														
<211	<b>&gt;</b> (	663														
<212	?> 1	ONA														
<213	3> i	Acrop	ora	aspe	era											
<220	)>															
<221	l> (	CDS									•					
<222	2>	(1)	. (663	3)												
<400	)> :	99														
		gtg Val														48
_	_	aat Asn						_	_		-				_	96
		gag Glu 35														144
	_	cca Pro		-		_					-	-	-			192
		cca Pro														240
cag Gln	tca Ser	ttc Phe	ccg Pro	gag Glu 85	gga Gly	ttt Phe	aca Thr	tgg Trp	gag Glu 90	agg Arg	atc Ile	atg Met	aac Asn	ttt Phe 95	gaa Glu	288
-		gca Ala		-		-	_		_		_					336
		atc Ile 115														384
		gtt Val	_	_	_	_		_			-					432

## 105/234

cgt ctc ttt gca Arg Leu Phe Ala 145	cga gat gga Arg Asp Gly 1 150	atg ctg ata gga Met Leu Ile Gly 155	aac aac ttt atg Asn Asn Phe Met	gct 480 Ala 160
ctg aag tta gaa Leu Lys Leu Glu				
tac aag gca aag Tyr Lys Ala Lys 180	aag cct gtg Lys Pro Val	aag atg cca ggg Lys Met Pro Gly 185	tat cat tat gtt Tyr His Tyr Val 190	gac 576 Asp
	Val Thr Asn		tac act tcc gtt Tyr Thr Ser Val 205	
cag tgt gaa att Gln Cys Glu Ile 210				663
<210> 100				
<211> 221				
<212> PRT				
<213> Acropora	aspera			
-	•			
<400> 100				
Mot Com Val Il-				
	Ala Thr Gin 5	Met Thr Tyr Lys 10	Val Tyr Met Ser 15	Gly
	5	10	15	_
Thr Val Asn Gly 20  Pro Tyr Glu Gly	5 His Tyr Phe Glu Gln Thr	10 Glu Val Glu Gly 25	Asp Gly Lys Gly 30  Val Thr Lys Gly	Lys
Thr Val Asn Gly 20 Pro Tyr Glu Gly 35	5 His Tyr Phe Glu Gln Thr	Glu Val Glu Gly 25 Val Arg Leu Ala 40	Asp Gly Lys Gly 30  Val Thr Lys Gly	Lys
Thr Val Asn Gly 20  Pro Tyr Glu Gly 35  Pro Leu Pro Phe	5 His Tyr Phe Glu Gln Thr Ala Trp Asp 55	Glu Val Glu Gly 25  Val Arg Leu Ala 40  Ile Leu Ser Pro	Asp Gly Lys Gly 30  Val Thr Lys Gly 45  Gln Cys Gln Tyr 60	Lys Gly Gly
Thr Val Asn Gly 20  Pro Tyr Glu Gly 35  Pro Leu Pro Phe 50  Ser Ile Pro Phe 65  Gln Ser Phe Pro	5 His Tyr Phe Glu Gln Thr Ala Trp Asp 55 Thr Lys Tyr 70	Glu Val Glu Gly 25  Val Arg Leu Ala 40  Ile Leu Ser Pro  Pro Glu Asp Ile 75	Asp Gly Lys Gly 30  Val Thr Lys Gly 45  Gln Cys Gln Tyr 60  Pro Asp Tyr Val	Lys Gly Gly Lys 80

-		106	/234	
Cys Phe Ile Tyı 115	: His Val Lys	Phe Ser Gly	Leu Asn Phe Pro 125	Pro Asn
Gly Pro Val Met 130	: Gln Lys Lys 135		Trp Glu Pro His 140	Ser Glu
Arg Leu Phe Ala 145	a Arg Asp Gly 150	Met Leu Ile	Gly Asn Asn Phe 155	Met Ala 160
Leu Lys Leu Gli	n Gly Gly Gly 165	His Tyr Leu 170	Cys Glu Phe Lys	Thr Thr 175
Tyr Lys Ala Lys 180		Lys Met Pro 185	Gly Tyr His Tyr 190	Val Asp
Arg Lys Leu Asp 195	o Val Thr Asn	His Asn Lys 200	Asp Tyr Thr Ser 205	Val Glu
Gln Cys Glu Ile 210	e Ser Ile Thr 215		Val Val Ala 220	
<210> 101				
<211> 663				
<212> DNA				
<213> Acropora	aspera			
<220>				
<221> CDS				
<222> (1)(66	53)			
			aag gtt tat atg Lys Val Tyr Met	
acg gtc aat gga Thr Val Asn Gly 20	cac tac ttt His Tyr Phe	gag gtc gaa Glu Val Glu 25	ggc gat gga aaa Gly Asp Gly Lys 30	gga aag 96 Gly Lys

cct tac gag ggg gag cag acg gta agg ctg gct gtc acc aag ggc gga Pro Tyr Glu Gly Glu Gln Thr Val Arg Leu Ala Val Thr Lys Gly Gly 35 40 45

144

### 107/234

cct Pro	ctg Leu 50	cca Pro	ttt Phe	gct Ala	tgg Trp	gat Asp 55	att Ile	tta Leu	tca Ser	cca Pro	cag G1n 60	tgt Cys	cag Gln	tac Tyr	gga Gly	192
agc Ser 65	ata Ile	cca Pro	ttc Phe	acc Thr	aag Lys 70	tac Tyr	cct Pro	gaa Glu	gac Asp	atc Ile 75	cct Pro	gac Asp	tat Tyr	gta Val	aag Lys 80	240
cag Gln	tca Ser	ttc Phe	ccg Pro	gag Glu 85	gga Gly	ttt Phe	aca Thr	tgg Trp	gag Glu 90	agg Arg	atc Ile	atg Met	aac Asn	ttt Phe 95	gaa Glu	288
gat Asp	ggt Gly	gca Ala	gtg Val 100	tgt Cys	act Thr	gtc Val	agc Ser	aat Asn 105	gat Asp	tcc Ser	agc Ser	atc Ile	caa Gln 110	ggc Gly	aac Asn	336
tgt Cys	ttc Phe	atc Ile 115	tac Tyr	cat His	gtc Val	aag Lys	ttc Phe 120	tct Ser	ggt Gly	ttg Leu	aac Asn	ttt Phe 125	cct Pro	ccc Pro	aat Asn	384
gga Gly	cct Pro 130	gtt Val	atg Met	cag Gln	aag Lys	aag Lys 135	aca Thr	cag Gln	ggc Gly	tgg Trp	gaa Glu 140	ccc Pro	cac His	tct Ser	gag Glu	432
cgt Arg 145	ctc Leu	ttt Phe	gca Ala	cga Arg	gat Asp 150	gga Gly	atg Met	ctg Leu	ata Ile	gga Gly 155	aac Asn	aac Asn	ttt Phe	atg Met	gct Ala 160	480
						ggt Gly										528
tac Tyr	aag Lys	gca Ala	aag Lys 180	aag Lys	cct Pro	gtg Val	aag Lys	atg Met 185	cca Pro	Gly ggg	tat Tyr	cat His	tat Tyr 190	gtt Val	gac Asp	576
						aat Asn										624
						aca Thr 215										663

<210> 102

<211> 221

<212> PRT

<213> Acropora aspera

<400> 102

Met Ser Val Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Ser Gly 1 5 10 15

#### 108/234

Thr Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys 20 25 30

Pro Tyr Glu Gly Glu Gln Thr Val Arg Leu Ala Val Thr Lys Gly Gly 35 40 45

Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Cys Gln Tyr Gly 50 55 60

Ser Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys 65 70 75 80

Gln Ser Phe Pro Glu Gly Phe Thr Trp Glu Arg Ile Met Asn Phe Glu 85 90 95

Asp Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn 100 105 110

Cys Phe Ile Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn 115 120 125

Gly Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro His Ser Glu 130 135 140

Arg Leu Phe Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala 145 150 155 160

Leu Lys Leu Glu Gly Gly Gly His Tyr Leu Cys Glu Phe Lys Thr Thr 165 170 175

Tyr Lys Ala Lys Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Asp 180 185 190

Arg Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu 195 200 205

Gln Cys Glu Ile Ser Ile Thr Arg Lys Pro Val Val Ala 210 215 220

<210> 103

<211> 663

<212> DNA

<213> Acanthastrea sp.

<220>	
<221> CDS	
<222> (1)(663)	
<400> $103$ atg agt gtg atc gct aca caa atg acc tac aag gtt tat atg tca ggc Met Ser Val Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Ser Gly $1$ $5$ $10$ $15$	48
acg gcc aat gga cac tac ttt gag gtt gaa ggc gat gga aaa gga aag Thr Ala Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys 20 25 30	96
cct tac gaa ggg gag cag acg gta agg ctc att gtc aca aag ggc gga Pro Tyr Glu Gly Glu Gln Thr Val Arg Leu Ile Val Thr Lys Gly Gly 35 40 45	44
cct ctg cca ttt gct tga gat att tta tca cca cag tat cag tac gga Pro Leu Pro Phe Ala Asp Ile Leu Ser Pro Gln Tyr Gln Tyr Gly 50 55 60	92
age ata cca tte ace aag tae cet gaa gae ate eet gae tat gta aag Ser Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys 65 70 75	40
cag tca ttc ccg gaa gga tat aca tgg gag agg atc atg aac ttt gaa Gln Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ile Met Asn Phe Glu 80 85 90 95	88
gat ggt gca gtg tgt act gtc agc aat gat tcc agc atc caa ggc aac Asp Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn 100 105 110	36
tgt ttc atc tac cat gtc aag ttc tct ggt ttg aac ttt cct ccc aat  Cys Phe Ile Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn  115 120 125	84
gga cct gtg atg cag aag aag aca cag ggc tgg gaa ccc aac act gag Gly Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu 130 135 140	32
cgt ctc ttt gca cga gat gga atg ctg ata gga aac aac ttt atg gct Arg Leu Phe Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala 145 150 155	80
ctg aag tta gaa ggc ggt cac tat ttg tgt gaa ttc aaa tct act Leu Lys Leu Glu Gly Gly His Tyr Leu Cys Glu Phe Lys Ser Thr 160 165 170 175	28
tac aag gca aag aag cct gtg aag atg cca ggg tat cac tat gtt gac Tyr Lys Ala Lys Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Asp 180 185 190	76

110/234

cgc aaa ctg gat gta acc aat cac aac aag gat tac act tcc gtt gag 624 Arg Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu 195 200

cag tgt gaa att tcc att gca cgc aaa cct gtg gtc gcc 663 Gln Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala

<210> 104

<211> 53

<212>

<213> Acanthastrea sp.

<400> 104

Met Ser Val Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Ser Gly

Thr Ala Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys

Pro Tyr Glu Gly Glu Gln Thr Val Arg Leu Ile Val Thr Lys Gly Gly 40

Pro Leu Pro Phe Ala 50

<210> 105

<211> 167

<212> PRT

<213> Acanthastrea sp.

<400> 105

Asp Ile Leu Ser Pro Gln Tyr Gln Tyr Gly Ser Ile Pro Phe Thr Lys 5 10

Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys Gln Ser Phe Pro Glu Gly

Tyr Thr Trp Glu Arg Ile Met Asn Phe Glu Asp Gly Ala Val Cys Thr

# 111/234

									111	1/234						
Val	Ser 50	Asn	Asp	Ser	Ser	Ile 55	Gln	Gly	Asn	Cys	Phe 60	Ile	Tyr	His	Val	
Lys 65	Phe	Ser	Gly	Leu	Asn 70	Phe	Pro	Pro	Asn	Gly 75	Pro	Val	Met	Gln	Lys 80	
Lys	Thr	Gln	Gly	Trp 85	Glu	Pro	Asn	Thr	Glu 90	Arg	Leu	Phe	Ala	Arg 95	Asp	
Gly	Met	Leu	Ile 100	Gly	Asn	Asn	Phe	Met 105	Ala	Leu	Lys	Leu	Glu 110	Gly	Gly	
Gly	His	Tyr 115	Leu	Cys	Glu	Phe	Lys 120	Ser	Thr	Tyr	Lys	Ala 125	Lys	Lys	Pro	
Val	Lys 130		Pro	Gly	Tyr	His 135	Tyr	Val	Asp	Arg	Lys 140	Leu	Asp	Val	Thr	
Asn 145		Asn	Lys	Asp	Tyr 150	Thr	Ser	Val	Glu	Gln 155	Cys	Glu	Ile	Ser	Ile 160	
Ala	Arg	Lys	Pro	Val 165	Val	Ala										
<21	0>	106														
<21	1>	663														
<21	2>	DNA														
<21	3> /	Acant	thast	trea	sp.											
<22	0>															
<22	1>	CDS														
<22	2>	(1)	. (663	3)												
atg		gtg											atg Met			48
													aaa Lys 30			96

.

WO 02/070703	PCT/GB02/00928

112/234

	tac Tyr														144
	ctg Leu 50				tga			tta Leu							192
-	ata Ile 65				_			-	-			_	_		240
	tca Ser														288
	ggt Gly														336
_	ttc Phe				-	-				_					384
	cct Pro														432
-	ctc Leu 145		-	_	-		_	_						_	480
-	aag Lys		_						_	-	-				528
	aag Lys														576
_	aaa Lys	_	-	_					_	_			_		624
-	tgt Cys	-				-	_				_	_			663

<210> 107

<211> 53

<212> PRT

<213> Acanthastrea sp.

<400> 107

#### 113/234

Met Ser Val Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Ser Gly 1 5 10 15

Thr Ala Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys 20 25 30

Pro Tyr Glu Gly Glu Gln Thr Val Arg Leu Ile Val Thr Lys Gly Gly 35  $\phantom{\bigg|}40\phantom{\bigg|}$ 

Pro Leu Pro Phe Ala 50

<210> 108

<211> 167

<212> PRT

<213> Acanthastrea sp.

<400> 108

Asp Ile Leu Ser Pro Gln Tyr Gln Tyr Gly Ser Ile Pro Phe Thr Lys
1 10 15

Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys Gln Ser Phe Pro Glu Gly 20 . 25 30

Tyr Thr Trp Glu Arg Ile Met Asn Phe Glu Asp Gly Ala Val Cys Thr 35 40 45

Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys Phe Ile Tyr His Val 50 55 60

Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly Pro Val Met Gln Lys 65 70 75 80

Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu Arg Leu Phe Ala Arg Asp 85 90 95

Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu Lys Leu Glu Gly Gly 100 105 110

Gly His Tyr Leu Cys Glu Phe Lys Ser Thr Tyr Lys Ala Lys Lys Pro 115 120 125

#### 114/234

Val Lys Met Pro Gly Tyr His Tyr Val Asp Arg Lys Leu Asp Val Thr 130 135 140

Asn His Asn Lys Asp Tyr Thr Ser Val Glu Gln Cys Glu Ile Ser Ile 145 150 155 160

Ala Arg Lys Pro Val Val Ala 165

<210> 109

<211> 663

<212> DNA

<213> Acanthastrea sp.

<220>

<221> CDS

<222> (1)..(663)

<400> 109

atg	agt	gtg	atc	gct	aca	caa	atg	acc	tac	aag	gtt	tat	atg	tca	ggc	48
Met	Ser	Val	I1e	Ala	Thr	Gln	Met	Thr	Tyr	Lys	Val	Tyr	Met	Ser	Gly	
1				5					10	-		•		15		

acg gcc aat gga cac tac ttt gag gtt gaa ggc gat gga aaa gga aag 96
Thr Ala Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys
20 25 30

cct tac gaa ggg gag cag acg gta agg ctc act gtc aca aag ggc gga 144
Pro Tyr Glu Gly Glu Gln Thr Val Arg Leu Thr Val Thr Lys Gly Gly
35

agc ata cca ttc acc aag tac cct gaa gac atc cct gac tat gta aag

Ser Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys

75

80

cag tca ttc ccg gaa gga tat aca tgg gag agg atc atg aac ttt gaa 288 Gln Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ile Met Asn Phe Glu

gat ggt gca gtg tgt act gtc agc aat gat tcc agc atc caa ggc aac
Asp Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn
100 105 110

		1:	15/234	
tgt ttc atc c Cys Phe Ile H 115	ac cat gtc a is His Val L	ag ttc tct gg ys Phe Ser G1 120	t ttg aac ttt cct y Leu Asn Phe Pro 125	ccc aat 384 Pro Asn
gga cct gtg a Gly Pro Val M 130	et Gln Lys L	ag aca cag gg ys Thr Gln Gl 35	c tgg gaa ccc aac y Trp Glu Pro Asn 140	act gag 432 Thr Glu
cgt ctc ttt g Arg Leu Phe A 145	ca cga gat g la Arg Asp G 150	ga atg ctg at ly Met Leu Il	a gga aac aac ttt e Gly Asn Asn Phe 155	atg gct 480 Met Ala 160
ctg aag tta g Leu Lys Leu G	aa gga ggc gg lu Gly Gly G 165	gt cac tat tt ly His Tyr Le 17	g tgt gaa ttc aaa u Cys Glu Phe Lys O	tct act 528 Ser Thr 175
Tyr Lys Ala L	ag aag cct g ys Lys Pro Va 80	tg aag atg cc al Lys Met Pro 185	a ggg tat cac tat o Gly Tyr His Tyr 190	gtt gac 576 Val Asp
cgc aaa ctg ga Arg Lys Leu A 195	at gta acc aa sp Val Thr As	at cac aac aa sn His Asn Ly 200	g gat tac act tcc s Asp Tyr Thr Ser 205	gtt gag 624 Val Glu
cag tgt gaa a Gln Cys Glu I 210	le Ser Ile A	ca cgc aaa cc la Arg Lys Pro 15	t gtg gtc gcc o Val Val Ala 220	663
<210> 110				
<211> 221				
<212> PRT				
<213> Acantha	astrea sp.			
<400> 110				
Met Ser Val II	le Ala Thr Gi 5	In Met Thr Ty: · 10	r Lys Val Tyr Met	Ser Gly 15
Thr Ala Asn G		ne Glu Val Glo 25	a Gly Asp Gly Lys	Gly Lys
Pro Tyr Glu G	ly Glu Gln Tì	nr Val Arg Le 40	Thr Val Thr Lys 45	Gly Gly

Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Tyr Gln Tyr Gly

Ser Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys 65 70 75 80

50

	W	O 02/0	07070	3												PCT/GB02/00
									110	5/234						•
Gln	Ser	Phe	Pro	Glu 85	Gly	Tyr	Thr	Tŕp	Glu 90	Arg	Ile	Met	Asn	Phe 95	Glu	
Asp	Gly	Ala	Val 100	Cys	Thr	Val	Ser	Asn 105	Asp	Ser	Ser	Ile	Gln 110	Gly	Asn	
Cys	Phe	Ile 115	His	His	Val	Lys	Phe 120	Ser	Gly	Leu	Asn	Phe 125	Pro	Pro	Asn	
Gly	Pro 130	Val	Met	Gln	Lys	Lys 135	Thr	Gln	Gly	Trp	Glu 140	Pro	Asn	Thr	Glu	
Arg 145	Leu	Phe	Ala	Arg	Asp 150	Gly	Met	Leu	Ile	Gly 155	Asn	Asn	Phe	Met	Ala 160	
Leu	Lys	Leu	Glu	Gly 165	Gly	Gly	His	Tyr	Leu 170	Cys	Glu	Phe	Lys	Ser 175	Thr	
Tyr	Lys	Ala	Lуs 180	Lys	Pro	Val	Lys	Met 185	Pro	Gly	Tyr	His	Tyr 190	Val	Asp	
Arg	Lys	Leu 195	Asp	Val	Thr	Asn	His 200	Asn	Lys	Asp	Tyr	Thr 205	Ser	Val	Glu	
Gln	Cys 210	Glu	Ile	Ser	Ile	Ala 215	Arg	Lys	Pro	Val	Val 220	Ala				
<210	O> :	111														
<21	1>	659														
<212	2> 1	DNA														
<213	3> 7	Acant	thas	trea	sp.											
<400 agt		111 tcg (	ctac	acaa	at ga	accta	acaa	g gt	ttat	atgt	cag	gcac	ggt	caat	ggaca	ac 60
tact	tttg	agg 1	tcga	aggc	ga t	ggaa	aagg	a aa	gcct	acga	ggg	ggag	cag	acgg	taaaq	gc 120

agtgtgatcg ctacacaaat gacctacaag gtttatatgt caggcacggt caatggacac 60
tactttgagg tcgaaggcga tggaaaagga aagcctacga gggggagcag acggtaaagc 120
tcactgtcac caagggcgga cctctgccat ttgcttggga tattttatca ccacagtgtc 180
agtacggaaa cataccattc accaagtacc ctgaagacgt ccctgactat gtaaagcagt 240
cattcccgga gggatttaca tgggagagga tcatgaactt tgaagatggt gcagtgtgta 300
ctgtcagcaa tgattccagc atccaaggca actgtttcac ctaccatgtc aagttctctg 360
gtttgaactt tcctcccaat ggacctgtga tgcagaagga gacacagggc tgggaacccc 420

actctgagcg tetet	ttgca cggggtggaa	a tgctgatagg a	aaacaacttt gtggctctg	a 480
agttagaagg aggcg	gtcac tatttgtgtg	g gattcaaaac t	tacttacaag gcaaagaaa	c 540
ctgtgaagat gccag	ggtat cattatgtt	g accgcaaact <u>c</u>	ggatgtaacc aatcacaac	a 600
aggattacat ttccg	ttgag cagtgtgaaa	a tttccattgc a	acgcaaacct gtggtcgcc	659
<210> 112				
<211> 663				
<212> DNA				
<213> Caulastre	a sn			
oudiabele.	a sp.			
<220>				
<221> CDS				
<222> (1)(663)				
<400> 112				
Met Ser Val Ile A	Ala Thr Gln Met	Thr Tyr Lys V	gtt tat atg tca ggc Val Tyr Met Ser Gly	48
	5	10	15	
			gat gga aaa gga aag Asp Gly Lys Gly Lys 30	96
	nag cag acg gta		gtc acc aag ggc gga	1.4.4
Pro Tyr Glu Gly (	Glu Gln Thr Val	Lys Leu Thr V	/al Thr Lys Gly Gly 45	144
		tta tca cca c	cag tgt cag tac gga	192
		Leu Ser Pro G	Gln Cys Gln Tyr Gly	132
aat ata cca ttc a	acc aag tac cct	gaa gac gtc c	ect gac tat gta aag	240
			Pro Asp Tyr Val Lys 80	
egg tea tte eeg e	gag gga ttt aca	tgg gag agg a	atc atg aac ttt gaa	288
	Glu Gly Phe Thr 85	Trp Glu Arg I 90	lle Met Asn Phe Glu 95	
gat ggt gca gtg t	tgt act gtc agc	aat gat too a	agc atc caa ggc aac	336
100	cys int val ser	105	Ser Ile Gln Gly Asn 110	
			aac ttt cct ccc aat Asn Phe Pro Pro Asn	384
115	120		125	

WO 02/070703 PCT/GB02/00928 118/234 gga cet gtg atg cag aag aag aca cag gge tgg gaa cee cae tet gag 432 Gly Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro His Ser Glu 130 135 cgt ctc ttt gca cgg ggt gga atg ctg ata gga aac aac ttt atg qct 480 Arg Leu Phe Ala Arg Gly Gly Met Leu Ile Gly Asn Asn Phe Met Ala 155 145 ctg aag tta gaa gga ggc ggt cac tat ttg tgt gga ttc aaa act act 528 Leu Lys Leu Glu Gly Gly Gly His Tyr Leu Cys Gly Phe Lys Thr Thr tac aag gca aag aag cct gtg aag atg cca ggg tat cat tat gtt gac 576 Tyr Lys Ala Lys Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Asp 180 185 190 cgc aaa ctg gat gta acc aat cac aac aag gat tac att tcc gtt gag 624 Arg Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Ile Ser Val Glu 200 cag tgt gaa att tcc att gca cgc aaa cct gtg gtc gcc 663 Gln Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 210 215 <210> 113 <211> 221

<212> PRT

<213> Caulastrea sp.

<400> 113

Met Ser Val Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Ser Gly 1 5 10 15

Thr Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys 20 25 30

Pro Tyr Glu Gly Glu Gln Thr Val Lys Leu Thr Val Thr Lys Gly Gly 35 40 45

Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Cys Gln Tyr Gly
50 55 60

Asn Ile Pro Phe Thr Lys Tyr Pro Glu Asp Val Pro Asp Tyr Val Lys 65 70 75 80

Arg Ser Phe Pro Glu Gly Phe Thr Trp Glu Arg Ile Met Asn Phe Glu 85 90 95

#### 119/234

Asp Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn 100 105 110

Cys Phe Thr Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn 115 120 125

Gly Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro His Ser Glu 130 135 140

Arg Leu Phe Ala Arg Gly Gly Met Leu Ile Gly Asn Asn Phe Met Ala 145 150 155 160

Leu Lys Leu Glu Gly Gly Gly His Tyr Leu Cys Gly Phe Lys Thr Thr 165 . 170 . 175

Tyr Lys Ala Lys Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Asp 180 185 190

Arg Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Ile Ser Val Glu 195  $\phantom{\bigg|}200\phantom{\bigg|}205\phantom{\bigg|}$ 

Gln Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 210 215 220

<210> 114

<211> 663

<212> DNA

<213> Caulastrea sp.

<220>

<221> CDS

<222> (1)..(663)

<400> 114

atg agt gtg atc gct aca caa atg acc tac aag gtt tat atg tcg ggc Met Ser Val Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Ser Gly 1 5 10 15

acg gtc aat gga cac tac ttt gag gtc gaa ggc gat gga aaa gga aag 96
Thr Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys
20 25 30

48

WO 02/070703		PCT/GB02/00928
	120/234	

									120	)/234						
				gag Glu												144
				gct Ala												192
				acc Thr												240
cag Gln	tca Ser	ttc Phe	ccg Pro	gag Glu 85	gga Gly	ttt Phe	aca Thr	tgg Trp	gag Glu 90	agg Arg	atc Ile	atg Met	aac Asn	ttt Phe 95	gaa Glu	288
		-		tgt Cys		-	_		-		_					336
				cat His												384
				cag Gln												432
_			-	cgg Arg			_	_						_	-	480
_	_		_	gga Gly 165					-	_						528
				aag Lys												576
Arg	Lys	Leu 195	Asp	gta Val	Thr	Asn	His 200	Asn	Lys	Asp	Tyr	11e 205				624
				tcc Ser			-				-	-				663

<210> 115

<211> 221

<212> PRT

<213> Caulastrea sp.

<400> 115

#### 121/234

Met Ser Val Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Ser Gly 1 5 10 15 .

Thr Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys 20 25 . 30

Pro Tyr Glu Gly Glu Gln Thr Val Lys Leu Thr Val Thr Lys Gly Gly 35 40 45

Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Cys Gln Tyr Gly 50 55

Asn Ile Pro Phe Thr Lys Tyr Pro Glu Asp Val Pro Asp Tyr Val Lys 65 70 75 80

Gln Ser Phe Pro Glu Gly Phe Thr Trp Glu Arg Ile Met Asn Phe Glu 85 90 95

Asp Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn 100 105 110

Cys Phe Thr Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn 115 120 125

Gly Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro His Ser Glu 130 135 140

Arg Leu Leu Ala Arg Gly Gly Met Leu Ile Gly Asn Asn Phe Met Ala 145 150 155 160

Leu Lys Leu Glu Gly Gly Gly His Tyr Leu Cys Gly Phe Lys Thr Thr 165 170 175

Tyr Lys Ala Lys Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Asp 180 185 190

Arg Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Ile Ser Val Glu 195 200 205

Gln Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 210 215 220

<210> 116

<211> 660

# WO 02/070703 PCT/GB02/00928 122/234

<212	2> 1	DNA														
<213	3> (	Caul	astr	ea s	p.											
<220	)>															
<221	L> (	CDS														
<222	2>	(1).	. (66	0)												
<400 agt Ser 1	gtg	116 atc Ile	gct Ala	aca Thr 5	caa Gln	atg Met	acc Thr	tac Tyr	aag Lys 10	gtt Val	tat Tyr	atg Met	tca Ser	ggc Gly 15	acg Thr	48
gtc Val	aat Asn	gga Gly	cac His 20	tac Tyr	ttt Phe	gag Glu	gtc Val	gaa Glu 25	ggc Gly	gat Asp	gga Gly	aaa Lys	gga Gly 30	aag Lys	cct Pro	96
tac Tyr	gag Glu	ggg Gly 35	gag Glu	cag Gln	acg Thr	gta Val	aag Lys 40	ctc Leu	act Thr	gtc Val	acc Thr	aag Lys 45	ggc Gly	gga Gly	cct Pro	144
ctg Leu	cca Pro 50	ttt Phe	gct A1a	tgg Trp	gat Asp	att Ile 55	tta Leu	tca Ser	cca Pro	cag Gln	tgt Cys 60	cag Gln	tac Tyr	gga Gly	aac Asn	192
ata Ile 65	cca Pro	ttc Phe	acc Thr	aag Lys	tac Tyr 70	cct Pro	gaa Glu	gac Asp	gtc Val	cct Pro 75	gac Asp	tat Tyr	gta Val	aag Lys	cag Gln 80	240
				gga Gly 85												288
				act Thr												336
ttc Phe	acc Thr	tac Tyr 115	cat His	gtc Val	aag Lys	ttc Phe	tct Ser 120	ggt Gly	ttg Leu	aac Asn	ttt Phe	cct Pro 125	ccc Pro	aat Asn	gga Gly	384
				aag Lys												432
				ggt Gly												480
				ggc Gly 165												528

WO 02/070703		PCT/GB02/00928
	123/234	

aag gca aag aag ctt gtg aag atg cca ggg tat cat tat gtt gac cgc
Lys Ala Lys Lys Leu Val Lys Met Pro Gly Tyr His Tyr Val Asp Arg
180 185 190

aaa ctg gat gta acc aat cac aac aag gat tac att tcc gtt gag cag
Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Ile Ser Val Glu Gln
195 200 205

tgt gaa att tcc att gca cgc aaa cct gtg gtc gcc
Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala

<210> 117

<211> 220

<212> PRT

<213> Caulastrea sp.

<400> 117

Ser Val Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr 1  $\phantom{\bigg|}$  5  $\phantom{\bigg|}$  10  $\phantom{\bigg|}$  15

Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys Pro 20 25 30

Tyr Glu Gly Glu Gln Thr Val Lys Leu Thr Val Thr Lys Gly Gly Pro 35 40 45

Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Cys Gln Tyr Gly Asn 50 55 60

Ile Pro Phe Thr Lys Tyr Pro Glu Asp Val Pro Asp Tyr Val Lys Gln 70 75 80

Ser Phe Pro Glu Gly Phe Thr Trp Glu Arg Ile Met Asn Phe Glu Asp 85 90 95

Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys 100 105 110

Phe Thr Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly 115 120 125

Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro His Ser Glu Arg 130 135 140

### 124/234

145		ly Met Leu 50		sn Asn Phe Met 55	Ala Leu 160
Lys Leu Glu	Gly Gly G 165	ly His Tyr	Leu Cys Gl	ly Phe Lys Th:	Thr Tyr 175
	Lys Leu Va 180	al Lys Met	Pro Gly Ty 185	yr His Tyr Va 190	
Lys Leu Asp 195	Val Thr A	sn His Asn 200		yr Ile Ser Va 205	l Glu Gln
Cys Glu Ile 210	Ser Ile Al	la Arg Lys 215	Pro Val Va	al Ala 220	
<210> 118					
<211> 660					
<212> DNA					
<213> Caula	strea sp.				
<220>					
<221> CDS					
<221> CDS <222> (1)	(660)				
<222> (1) <400> 118		aa atg acc	tac aad ot		n ggg 20g 48
<222> (1) <400> 118 agt gtg atc	get aca ca	aa atg acc ln Met Thr	tac aag gt Tyr Lys Va 10	it tat atg tca al Tyr Met Sei	a ggc acg 48 c Gly Thr 15
<222> (1)  <400> 118 agt gtg atc Ser Val Ile 1  gtc aat gga Val Asn Gly	gct aca ca Ala Thr Gi 5 cac tac ti	ln Met Thr tt gag gtc	Tyr Lys Va 10 gaa ggc ga	et tat atg toa al Tyr Met Sen at gga aaa gga sp Gly Lys Gly 30	Gly Thr 15 a aag cct 96
<222> (1)  <400> 118 agt gtg atc Ser Val Ile 1 gtc aat gga Val Asn Gly tac gag ggg	gct aca ca Ala Thr G 5 cac tac t His Tyr P 20 gag cag ac	In Met Thr tt gag gtc ne Glu Val	Tyr Lys Va 10 gaa ggc ga Glu Gly As 25 ctc act gt	al Tyr Met Sen at gga aaa gga ap Gly Lys Gly	Gly Thr 15 a aag cct 96 7 Lys Pro
<222> (1)  <400> 118 agt gtg atc Ser Val Ile 1  gtc aat gga Val Asn Gly  tac gag ggg Tyr Glu Gly 35  ctg cca ttt	gct aca ca Ala Thr G: 5  cac tac to His Tyr Pi 20  gag cag ac Glu Gln Ti	In Met Thr  tt gag gtc  ne Glu Val  cg gta aag  nr Val Lys  40  at att tta	Tyr Lys Variation 10  gaa ggc ga Glu Gly As 25  ctc act gt Leu Thr Variation 10	al Tyr Met Sen at gga aaa gga sp Gly Lys Gly 30 cc acc aag gga al Thr Lys Gly	Gly Thr 15 a aag cct 96 7 Lys Pro c gga cct 144 7 Gly Pro c gga aac 192

WO 02/070703	PCT/GB02/00928
--------------	----------------

125/234 tca ttc ccg gag gga ttt aca tgg gag agg atc atg aac ttt gaa gat 288 Ser Phe Pro Glu Gly Phe Thr Trp Glu Arg Ile Met Asn Phe Glu Asp ggt gca gtg tgt act gtc agc aat gat tcc agc atc caa ggc aac tgt 336 Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys 100 105 ttc acc tac cat gtc aag ttc tct ggt ttg aac ttt cct ccc aat gga 384 Phe Thr Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly 120 cct gtg atg tag aag aag aca cag ggc tgg gaa ccc cac tct gag cgt 432 Pro Val Met Lys Lys Thr Gln Gly Trp Glu Pro His Ser Glu Arg 135 ctc ttt gca cgg ggt gga atg ctg ata gga aac aac ttt atg gct ctg 480 Leu Phe Ala Arg Gly Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu 155 aag tta gaa gga ggt ggt cac tat ttg tgt gaa ttc aaa act act tac 528 Lys Leu Glu Gly Gly His Tyr Leu Cys Glu Phe Lys Thr Thr Tyr 165 170 aag gca aag aag cct gtg aag atg cca ggg tat cat tat gtt gac cgc 576 Lys Ala Lys Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Asp Arg

180 185 aaa ctg gat gta acc aat cac aac aag gat tac att tcc gtt gag cag 624 Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Ile Ser Val Glu Gln 200

tgt gaa att tcc att gca cgc aaa cct gtg gtc gcc 660 Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 215

<210> 119

<211> 131

<212> PRT

<213> Caulastrea sp.

<400> 119

Ser Val Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr 5

Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys Pro

Tyr Glu Gly Glu Gln Thr Val Lys Leu Thr Val Thr Lys Gly Gly Pro 35 40

#### 126/234

Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Cys Gln Tyr Gly Asn 50 55

Ile Pro Phe Thr Lys Tyr Pro Glu Asp Val Pro Asp Tyr Val Lys Gln 65 70 75 80

Ser Phe Pro Glu Gly Phe Thr Trp Glu Arg Ile Met Asn Phe Glu Asp 85 90 95

Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys 100 105 110

Phe Thr Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly 115 120 125

Pro Val Met 130

<210> 120

<211> 88

<212> PRT

<213> Caulastrea sp.

<400> 120

Lys Lys Thr Gln Gly Trp Glu Pro His Ser Glu Arg Leu Phe Ala Arg 1  $\phantom{\bigg|}$  5  $\phantom{\bigg|}$  10  $\phantom{\bigg|}$  15

Gly Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu Lys Leu Glu Gly 20 25 30

Gly Gly His Tyr Leu Cys Glu Phe Lys Thr Thr Tyr Lys Ala Lys Lys 35 40 45

Pro Val Lys Met Pro Gly Tyr His Tyr Val Asp Arg Lys Leu Asp Val 50 55 60

Thr Asn His Asn Lys Asp Tyr Ile Ser Val Glu Gln Cys Glu Ile Ser 65 70 75 80

Ile Ala Arg Lys Pro Val Val Ala

### 127/234

<210> 121	
<211> 663	
<212> DNA	
<213> Acropora nobilis	
<220>	
<221> CDS	
<222> (1)(663)	
<pre>&lt;400&gt; 121 atg agt gtg atc gct aca caa atg acc tac aag gtt tat atg tca ggc</pre>	4.0
Met Ser Val Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Ser Gly	48
acg gtc aat gga cac tac ttt gag gtt gaa ggc gat gga aaa gga aag	96
Thr Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys 20 25 30	90
cct tac gaa ggg gag cag acg gta agg ctc act gtc aca aag ggc gga	144
Pro Tyr Glu Gly Glu Gln Thr Val Arg Leu Thr Val Thr Lys Gly Gly 35 40 45	144
cct ctg cca ttt gct tgg gat att tta tca cca cag tat cag tac gga	192
Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Tyr Gln Tyr Gly 50 55 60	132
ago ata coa tto aco aag tao cot gaa gao ato cot gao tat gta aag	240
Ser Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys 65 70 75 80	
cag toa tto cog gaa gga tat aca tgg gag agg ato atg aac ttt gaa	288
Gln Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ile Met Asn Phe Glu 85 90 95	
gat ggt gca gtg tgt act gtc agc aat gat tec agc atc caa ggc aac	336
Asp Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn 100 105 110	
tgt tte ate tac cat gte aag tte tet ggt ttg aac ttt eet eec aac	384
Cys Phe Ile Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn 115 120 125	
gga cct gtg atg Cag aag aag aca cag ggc tgg gaa ccc aac act gag	432
Gly Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu 130 135 140	
cgt ctc tta gca cga gat gga atg ctg ata gga aac aac ttt atg gct	480
Arg Leu Leu Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala 145 155 160	

WO 02/070703	PCT/GB02/00928
· · · · · · · · · · · · · · · · · · ·	

WO 02	/070703			PCT/GB02/06
			128/234	·
ctg aag tta Leu Lys Leu	gaa gga ggc Glu Gly Gly 165	ggt cac ta Gly His Ty	t ttg tgt gaa ttc r Leu Cys Glu Phe 170	aaa tct act 528 Lys Ser Thr 175
tac aag gca Tyr Lys Ala	a aag aag cct Lys Lys Pro 180	gtg aag at Val Lys Me 18	g cca ggg tat cac t Pro Gly Tyr His 5	tat gtt gac 576 Fyr Val Asp 190
cgc aaa cto Arg Lys Leu 195	Asp Val Thr	aat cac aa Asn His As 200	c aag gat tac act n Lys Asp Tyr Thr 205	tcc gtt gag 624 Ser Val Glu
			a cct gtg gtc gcc s Pro Val Val Ala 220	663
<210> 122				
<211> 221				
<212> PRT				
<213> Acro	pora nobilis			
<400> 122				
Met Ser Val	. Ile Ala Thr 5	Gln Met Th	r Tyr Lys Val Tyr 10	Met Ser Gly 15
Thr Val Asn	Gly His Tyr 20	Phe Glu Va 25	l Glu Gly Asp Gly	Lys Gly Lys 30
Pro Tyr Glu 35	Gly Glu Gln	Thr Val Ar	g Leu Thr Val Thr 45	Lys Gly Gly
Pro Leu Pro 50	Phe Ala Trp	Asp Ile Le	u Ser Pro Gln Tyr 60	Gln Tyr Gly
Ser Ile Pro	Phe Thr Lys	Tyr Pro Gl	u Asp Ile Pro Asp 75	fyr Val Lys 80
Gln Ser Phe	Pro Glu Gly 85	Tyr Thr Tr	p Glu Arg Ile Met 90	Asn Phe Glu 95
Asp Gly Ala	Val Cys Thr 100	Val Ser As:	n Asp Ser Ser Ile 5	Gln Gly Asn 110

Cys Phe Ile Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn 115 120 125

### 129/234

Gly Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu 130 135 140	
Arg Leu Leu Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala 145 150 155 160	
Leu Lys Leu Glu Gly Gly Gly His Tyr Leu Cys Glu Phe Lys Ser Thr 165 170 175	
Tyr Lys Ala Lys Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Asp 180 185 190	
Arg Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu 195 200 205	
Gln Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 210 215 220	
<210> 123	
<211> 663	
<212> DNA	
<213> Acropora nobilis	
<220>	
<221> CDS	
<222> (1)(663)	
<222> (1)(663) <400> 123	
<222> (1)(663) <400> 123	18
<222> (1)(663)  <400> 123 atg agt gtg atc gct aca caa atg acc tac aag gtt tat atg tca ggc Met Ser Val Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Ser Gly 1 5 10 15	18
<pre>&lt;222&gt; (1)(663)  &lt;400&gt; 123 atg agt gtg atc gct aca caa atg acc tac aag gtt tat atg tca ggc Met Ser Val Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Ser Gly 1</pre>	96

WO 02/070703		PCT/GB02/00928
•	130/234	

agc Ser 65	ata Ile	cca Pro	ttc Phe	acc Thr	aag Lys 70	tac Tyr	cct Pro	gaa Glu	gac Asp	atc Ile 75	cct Pro	gac Asp	tat Tyr	gta Val	aag Lys 80	240
tag	tca Ser	ttc Phe	ccg Pro	gag Glu	gga Gly 85	ttt Phe	aca Thr	tgg Trp	gac Asp	agg Arg 90	atc Ile	atg Met	gac Asp	ttt Phe	gaa Glu 95	288
gat Asp	ggt Gly	gca Ala	gtg Val	tgt Cys 100	acc Thr	gtc Val	agc Ser	aat Asn	gat Asp 105	tcc Ser	agc Ser	atc Ile	caa Gln	ggc Gly 110	aac Asn	336
tgt Cys	ttc Phe	atc Ile	tac Tyr 115	cat His	gtc Val	aag Lys	ttc Phe	tct Ser 120	ggt Gly	ttg Leu	aac Asn	ttt Phe	cct Pro 125	ccc Pro	aat Asn	384
gga Gly	cct Pro	gtt Val 130	atg Met	cag Gln	aag Lys	aag Lys	aca Thr 135	cag Gln	ggc Gly	tgg Trp	gaa Glu	ccc Pro 140	aac Asn	act Thr	gag Glu	432
					gat Asp											480
					ggt Gly 165											528
tac Tyr	aag Lys	gca Ala	aag Lys	aag Lys 180	cct Pro	gtg Val	aag Lys	atg Met	cca Pro 185	Gly ggg	tat Tyr	cac His	tat Tyr	gtt Val 190	gac Asp	576
cgc Arg	aaa Lys	ctg Leu	gat Asp 195	gta Val	acc Thr	aat Asn	cac His	aac Asn 200	aag Lys	gat Asp	tac Tyr	act Thr	tcc Ser 205	gtt Val	gag Glu	624
					att Ile											663
<210	)> ]	124														

<211> 80

<212> PRT

<213> Acropora nobilis

<400> 124

Met Ser Val Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Ser Gly 1 5 10 15

Thr Val Asn Gly His Tyr Leu Glu Val Glu Gly Asp Gly Lys Gly Lys 20 25 30

131/234

Pro Tyr Glu Gly Glu Gln Thr Val Arg Leu Thr Val Thr Lys Gly Gly 35 40

Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Tyr Gln Tyr Gly 50 55 60

Ser Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys 65 70 75 80

<210> 125

<211> 140

<212> PRT

<213> Acropora nobilis

<400> 125

Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys 20 25 30

Phe Ile Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly 35 40 45

Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu Arg 50 60

Leu Phe Ala Arg Asp Gly Met Leu Leu Gly Asn Asn Phe Met Ala Leu 65 70 75 80

Lys Leu Glu Gly Gly His Tyr Leu Cys Glu Phe Lys Thr Thr Tyr 85 90 95

Lys Ala Lys Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Asp Arg 100 105 110

Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu Gln 115 120 125

Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 130  $$135\$ 

### WO 02/070703 PCT/GB02/00928 132/234

<210> 126

€ <211> 660

<212> DNA

<213> Millepora sp.

<220>

<221> CDS

<222> (1)..(660)

	O>_ :															
													tca Ser			48
													gga Gly 30			96
													ggc Gly			144
_			_		-					_	_	_	tac Tyr		-	192
				-			_	_			_		gta Val	_	_	240
		_									_		ttt Phe	_		288
	_		_		_	-		_		_			ggc Gly 110		_	336
ttc Phe	acc Thr	tac Tyr 115	cat His	gtc Val	aag Lys	ttc Phe	tct Ser 120	ggt Gly	ttg Leu	aac Asn	ttt Phe	cct Pro 125	ccc Pro	aat Asn	gga Gly	384
		_	_	_	_		_			_			tct Ser		_	432
		_				_	_						atg Met	-	_	480

WO 02/070703	PCT/GB02/00928
W U 02/070703	FC1/GB02/00928

	***	0 02/0	70702	•					133	/234						101/0	1 <b>D</b> U2/U
aag Lys	tta Leu	gaa Glu	gga Gly	ggc Gly 165	ggt Gly	cac His	tat Tyr	ttg Leu	tgt Cys 170	gaa Glu	ttc Phe	aaa Lys	act Thr	act Thr 175	tac Tyr		528
aag Lys	gca Ala	aag Lys	aag Lys 180	cct Pro	gtg Val	aag Lys	atg Met	cca Pro 185	ggg Gly	tat Tyr	cat His	tat Tyr	gtt Val 190	gac Asp	cgc Arg		576
aaa Lys	ctg Leu	gat Asp 195	gta Val	acc Thr	aat Asn	cac His	aac Asn 200	aag Lys	gat Asp	tac Tyr	act Thr	tcc Ser 205	gtt Val	gag Glu	cag Gln		624
		att Ile			Ala												660
<210	)> :	127		•													
<211	i> :	220															
<212	2> 1	PRT															
<213	3> 1	Mille	epora	a sp	•												
<400	)> :	127															
Ser 1	Val	Ile	Ala	Thr 5	Gln	Met	Thr	Tyr	Lys 10	Val	Tyr	Met	Ser	Gly 15	Thr		
Val	Asp	Gly	His 20	Tyr	Phe	Glu	Val	G1u 25	Gly	Asp	Gly	Lys	Gly 30	Lys	Pro		
Туr	Glu	Gly 35	Glu	Gln	Thr	Val	Lys 40	Leu	Thr	Val	Thr	Lys 45	Gly	Gly	Pro		
Leu	Pro 50	Phe	Ala	Trp	Asp	Ile 55	Leu	Ser	Pro	Gln	Cys 60	Gln	Tyr	Gly	Ser		
Ile 65	Pro	Phe	Thr	Lys	Tyr 70	Pro	Glu	Asp	Ile	Pro 75	Asp	Tyr	Val	Lys	Gln 80		
Ser	Phe	Pro	Glu	Gly 85	Phe	Thr	Trp	Glu	Arg 90	Ile	Met	Asn	Phe	Glu 95	Asn		
Gly	Ala	Val	Cys 100	Thr	Val	Ser	Asn	Asp. 105	Ser	Ser	Ile	Gln	Gly 110	Asn	Cys		
Phe	Thr	Tyr 115	His	Val	Lys	Phe	Ser 120	Gly	Leu	Asn	Phe	Pro 125	Pro	Asn	Gly		

### 134/234

Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro His Ser Glu Arg 130 135 140	
Leu Phe Ala Arg Gly Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu 145 150 155 160	
Lys Leu Glu Gly Gly Gly His Tyr Leu Cys Glu Phe Lys Thr Thr Tyr 165 170 175	
Lys Ala Lys Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Asp Arg 180 185 190	
Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu Gln 195 200 205	
Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 210 215 220	
<210> 128	
<211> 663	
<212> DNA	
<213> Millepora sp.	
<220>	
<221> CDS	
<222> (1)(663)	
<400> 128	40
atg agt gtg atc gct aca caa atg acc tac aag gtt tat atg tca ggc Met Ser Val Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Ser Gly 1 10 15	48
acg gtc gat gga cac tac ttt gag gtc gaa ggc gat gga aaa gga aag Thr Val Asp Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys 20 25 30	96
cct tac gag ggg gag cag acg gta aag ctc act gtc acc aag ggc gga Pro Tyr Glu Gly Glu Gln Thr Val Lys Leu Thr Val Thr Lys Gly Gly 35 40 45	144
cct ctg cca ttt gct tgg gat att tta tca cca cag tgt cag tac gga Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Cys Gln Tyr Gly 50 55 60	192

WO 02/070703			PCT/GB02/00928
	405/034	•	

								•	135	/234						
agc Ser 65	ata Ile	cca Pro	ttc Phe	acc Thr	aag Lys 70	tac Tyr	cct Pro	gaa Glu	gac Asp	atc Ile 75	cct Pro	gac Asp	tat Tyr	gta Val	aag Lys 80	240
cag Gln	tca Ser	ttc Phe	ccg Pro	gag Glu 85	gga Gly	ttt Phe	aca Thr	tgg Trp	gag Glu 90	agg Arg	atc Ile	atg Met	aac Asn	ttt Phe 95	gaa Glu	288
gat Asp	ggt Gly	gca Ala	gtg Val 100	tgt Cys	act Thr	gtc Val	agc Ser	aat Asn 105	gat Asp	tcc Ser	agc Ser	atc Ile	caa Gln 110	ggc Gly	aac Asn	336
tgt Cys	ttc Phe	acc Thr 115	tac Tyr	cat His	gtc Val	aag Lys	ttc Phe 120	tct Ser	ggt Gly	ttg Leu	aac Asn	ttt Phe 125	cct Pro	ccc Pro	aat Asn	384
gga Gly	cct Pro 130	gtg Val	atg Met	cag Gln	aag Lys	aag Lys 135	aca Thr	cag Gln	ggc Gly	tgg Trp	gaa Glu 140	ccc Pro	cac His	tct Ser	gag Glu	432
								ctg Leu								480
ctg Leu	aag Lys	tta Leu	gaa Glu	gga Gly 165	ggc Gly	ggt Gly	cac His	tat Tyr	ttg Leu 170	tgt Cys	gaa Glu	ttc Phe	aaa Lys	act Thr 175	act Thr	528
tac Tyr	aag Lys	gca Ala	aag Lys 180	aag Lys	cct Pro	gtg Val	aag Lys	atg Met 185	cca Pro	Gly ggg	tat Tyr	cat His	tat Tyr 190	gtt Val	gac Asp	576
								aac Asn								624
_	_	_				-	_	aaa Lys			_	_				663
<210	)> ]	29														
<211	l> 2	221														
<212	?> I	PRT														

<213> Millepora sp.

<400> 129

Met Ser Val Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Ser Gly

Thr Val Asp Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys 20 25 30

136/234

Pro Tyr Glu Gly Glu Gln Thr Val Lys Leu Thr Val Thr Lys Gly Gly 35 40

Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Cys Gln Tyr Gly 50 55 60

Ser Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys 65 70 75 80

Gln Ser Phe Pro Glu Gly Phe Thr Trp Glu Arg Ile Met Asn Phe Glu 85 90 95

Asp Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn 100 105 110

Cys Phe Thr Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn 115 120 125

Gly Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro His Ser Glu 130 135 140

Arg Leu Phe Ala Arg Gly Gly Met Leu Ile Gly Asn Asn Phe Met Ala 145 150 155 160

Leu Lys Leu Glu Gly Gly Gly His Tyr Leu Cys Glu Phe Lys Thr Thr 165 170 175

Tyr Lys Ala Lys Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Asp 180 185 190

Arg Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu 195 200 205

Gln Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 210 215 220

<210> 130

<211> 663

<212> DNA

<213> Millepora sp.

<220>

### 137/234

<221> CDS

<222> (1)..(663)

<40		130															
atg Met 1	agt Ser	gtg Val	atc Ile	gct Ala 5	aca Thr	caa Gln	atg Met	acc Thr	tac Tyr 10	aag Lys	gtt Val	tat Tyr	atg Met	tca Ser 15	ggc		48
acg Thr	gtc Val	gat Asp	gga Gly 20	cac His	tac Tyr	ttt Phe	gag Glu	gtc Val 25	gaa Glu	ggc Gly	gat Asp	gga Gly	aaa Lys 30	gga Gly	aag Lys		96
cct Pro	tac Tyr	gag Glu 35	ggg Gly	gag Glu	cag Gln	acg Thr	gta Val 40	aag Lys	ctc Leu	act Thr	gtc Val	acc Thr 45	aag Lys	ggc Gly	gga Gly	1	L44
cct Pro	ctg Leu 50	cca Pro	ttt Phe	gct Ala	tgg Trp	gat Asp 55	att Ile	tta Leu	tca Ser	cca Pro	cag Gln 60	tgt Cys	cag Gln	tac Tyr	gga Gly	1	192
agc Ser 65	ata Ile	cca Pro	ttc Phe	acc Thr	aag Lys 70	tac Tyr	cct Pro	gaa Glu	gac Asp	atc Ile 75	cct Pro	gac 'Asp	tat Tyr	gta Val	aag Lys 80		240
	tca Ser															. 2	288
gat Asp	ggt Gly	gca Ala	gtg Val 100	tgt Cys	act Thr	gtc Val	agc Ser	aat Asn 105	ggt Gly	tcc Ser	agc Ser	atc Ile	caa Gln 110	ggc Gly	aac Asn	3	336
tgt Cys	ttc Phe	acc Thr 115	tac Tyr	cat His	gtc Val	aag Lys	ttc Phe 120	tct Ser	ggt Gly	ttg Leu	aac Asn	ttt Phe 125	cct Pro	ccc Pro	aat Asn	3	884
gga Gly	cct Pro 130	gtg Val	atg Met	cag Gln	aag Lys	aag Lys 135	aca Thr	cag Gln	ggc Gly	tgg Trp	gaa Glu 140	ccc Pro	cac His	tct Ser	gag Glu	4	32
cgt Arg 145	ctc Leu	ttt Phe	gca Ala	cgg Arg	ggt Gly 150	gga Gly	atg Met	ctg Leu	ata Ile	gga Gly 155	aac Asn	aac Asn	ttt Phe	atg Met	gct Ala 160	4	180
	aag Lys															5	28
tac Tyr	agg Arg	gca Ala	aag Lys 180	aag Lys	cct Pro	gtg Val	aag Lys	atg Met 185	cca Pro	Gly ggg	tat Tyr	cat His	tat Tyr 190	gtt Val	gac Asp	5	76
cgc Arg	aaa Lys	ctg Leu 195	gat Asp	gta Val	acc Thr	aat Asn	cac His 200	aac Asn	aag Lys	gat Asp	tac Tyr	act Thr 205	tcc Ser	gtt Val	gag Glu	6	24

#### 138/234

cag tgt gaa att tcc att gca cgc aaa cct gtg gtc gcc Gln Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 210 215 220

663

<210> 131

<211> 221

<212> PRT

<213> Millepora sp.

<400> 131

Met Ser Val Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Ser Gly
1 5 10 15

Thr Val Asp Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys 20 25 30

Pro Tyr Glu Gly Glu Gln Thr Val Lys Leu Thr Val Thr Lys Gly Gly 35 40 45

Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Cys Gln Tyr Gly 50 60

Ser Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys 65 70 75 80

Gln Ser Phe Pro Glu Gly Phe Thr Trp Glu Arg Ile Met Asn Phe Glu 85 90 95

Asp Gly Ala Val Cys Thr Val Ser Asn Gly Ser Ser Ile Gln Gly Asn 100 105 110

Cys Phe Thr Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn 115 120 125

Gly Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro His Ser Glu 130 135 140

Arg Leu Phe Ala Arg Gly Gly Met Leu Ile Gly Asn Asn Phe Met Ala 145 150 155 160

Leu Lys Leu Gly Gly Gly His Tyr Leu Cys Glu Phe Lys Thr Thr 165 170 175

#### 139/234

Tyr Arg Ala	Lys Lys	Pro V	al Lys	Met	Pro	Gly	Tyr	His	Tyr	Val	Asp
	180			185					190		-

Arg Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu 195 200 205

Gln Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 210 215 220

<210> 132

<211> 660

<212> DNA

<213> Millepora sp.

<220>

<221> CDS

<400> 132

<222> (1)..(660)

100

agt Ser 1	gtg Val	atc Ile	gct Ala	aca Thr 5	caa Gln	atg Met	acc Thr	tac Tyr	aag Lys 10	gtt Val	tat Tyr	atg Met	tca Ser	ggc Gly 15	acg Thr	48	
gtc Val	gat Asp	gga Gly	cac His 20	tac Tyr	ttt Phe	gag Glu	gtc Val	gaa Glu 25	ggc Gly	gat Asp	gga Gly	aaa Lys	gga Gly 30	aag Lys	cct Pro	96	
					acg Thr											144	
_			-		gat Asp					_	-	-			_	192	
					tac Tyr 70											240	
					ttt Phe											288	
		Val		Thr	gtc Val	Ser	Asn	Asp	Ser	Ser	Ile	Gln				336	

	WC	) 02/0	70703	i												PCT/GB02/009
									140	/234						
	acc Thr															384
cct Pro	gtg Val 130	atg Met	cag Gln	aag Lys	aag Lys	aca Thr 135	cag Gln	ggc Gly	tgg Trp	gaa Glu	ccc Pro 140	cac His	tct Ser	gag Glu	cgt Arg	432
ctc Leu 145	ttt Phe	gca Ala	cgg Arg	ggt Gly	gga Gly 150	atg Met	ctg Leu	ata Ile	gga Gly	aac Asn 155	aac Asn	ttt Phe	atg Met	gct Ala	ctg Leu 160	480
	tta Leu															528
	gca Ala															576
	ctg Leu															624
	gaa Glu 210															660
<210	<b>)</b> > ]	133														
<21.	1> 2	220														
<21	2> I	PRT														
<21	1 <8	Mille	epora	a sp	•											
<40	0> 1	133														
Ser	Val	Ile	Ala	Thr	Gln	Met	Thr	Tyr	Lys	Val	Tyr	Met	Ser	Gly	Thr	

Val Asp Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys Pro 20 30

Tyr Glu Gly Glu Gln Thr Val Lys Leu Thr Val Thr Lys Gly Gly Pro 35 40 45

Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Cys Gln Tyr Gly Ser 50

Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys Gln 70

#### 141/234

Ser Phe Pro Glu Gly Phe Thr Trp Glu Arg Ile Met Asn Phe Glu Asp 85 90 95

Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys 100 105 110

Phe Thr Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly 115 120 125

Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro His Ser Glu Arg 130 135 140

Leu Phe Ala Arg Gly Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu 145 150 155 160

Lys Leu Glu Gly Gly His Tyr Leu Cys Glu Phe Lys Thr Thr Tyr 165 170 175

Lys Ala Lys Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Asp Arg 180 185 190

Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu Gln 195 200 205

Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 210 215 220

<210> 134

<211> 663

<212> DNA

<213> Porites murrayensis

<220>

<221> CDS

<222> (1)..(663)

<400> 134

atg agt gtg atc gct aca caa atg acc tac aag gtt tat atg cca ggc Met Ser Val Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Pro Gly 1 5 10 15

WO 02/070703		PCT/GB02/00928
	142/234	

									172	.1234						
acg Thr	gtc Val	aat Asn	gga Gly 20	cac His	tac Tyr	ttt Phe	gag Glu	gtt Val 25	gaa Glu	ggc Gly	gat Asp	gga Gly	aaa Lys 30	gga Gly	aag Lys	96
cct Pro	tac Tyr	gag Glu 35	ggg Gly	gag Glu	cag Gln	acg Thr	gta Val 40	aag Lys	ctc Leu	act Thr	gtc Val	acc Thr 45	aag Lys	ggc Gly	gga Gly	144
cct Pro	ctg Leu 50	cca Pro	ttt Phe	gct Ala	tgg Trp	gat Asp 55	att Ile	cta Leu	tca Ser	cca Pro	cag G1n 60	agt Ser	cag Gln	tac Tyr	gga Gly	192
agc Ser 65	ata Ile	cca Pro	ttc Phe	acc Thr	aag Lys 70	tac Tyr	cct Pro	gaa Glu	gac Asp	atc Ile 75	cct Pro	gac Asp	tat Tyr	gta Val	aag Lys 80	240
cag Gln	tca Ser	ttc Phe	cct Pro	gag Glu 85	gga Gly	tat Tyr	aca Thr	tgg Trp	gag Glu 90	agg Arg	atc Ile	atg Met	aac Asn	ttc Phe 95	gaa Glu	288
gat Asp	ggt Gly	gca Ala	gtg Val 100	tgt Cys	act Thr	gtc Val	agc Ser	aat Asn 105	gat Asp	tcc Ser	agc Ser	atc Ile	caa Gln 110	ggt Gly	aac Asn	336
tgt Cys	ttc Phe	atc Ile 115	tac Tyr	aat Asn	gtc Val	aag Lys	ttc Phe 120	tct Ser	ggt Gly	ttg Leu	aac Asn	ttt Phe 125	cct Pro	ccc Pro	aat Asn	384
gga Gly	cct Pro 130	gtt Val	atg Met	caa Gln	aag Lys	aag Lys 135	aca Thr	cag Gln	ggc Gly	tgg Trp	gaa G1u 140	ccc Pro	aac Asn	act Thr	gag Glu	432
cgt Arg 145	ctt Leu	tat <b>T</b> yr	gca Ala	cga Arg	gat Asp 150	gga Gly	atg Met	ctg Leu	ata Ile	gga Gly 155	aac Asn	aac Asn	ttt Phe	atg Met	gct Ala 160	480
ctg Leu	aag Lys	ttg Leu	gaa Glu	gga Gly 165	ggt Gly	ggt Gly	cat His	tat Tyr	ttg Leu 170	tgt Cys	gaa Glu	ttc Phe	aaa Lys	tct Ser 175	act Thr	528
tac Tyr	aag Lys	gca Ala	aag Lys 180	aag Lys	cct Pro	gtg Val	atg Met	atg Met 185	cct Pro	gga Gly	tat Tyr	cac His	tat Tyr 190	gtt Val	gac Asp	576
cgc Arg	aaa Lys	ttg Leu 195	gat Asp	gta Val	acc Thr	aat Asn	cac His 200	aac Asn	aag Lys	gat Asp	tac Tyr	act Thr 205	tcc Ser	gtt Val	gag Glu	624
	tgt Cys 210															663

<210> 135

<211> 221

<212> PRT

<213> Porites murrayensis

PCT/GB02/00928 143/234

<400> 135

Met Ser Val Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Pro Gly 10 . 15

Thr Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys

Pro Tyr Glu Gly Glu Gln Thr Val Lys Leu Thr Val Thr Lys Gly Gly

Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Ser Gln Tyr Gly

Ser Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys 70 . 75 80

Gln Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ile Met Asn Phe Glu 85 90

Asp Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn

Cys Phe Ile Tyr Asn Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn 120

Gly Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu 130

Arg Leu Tyr Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala

Leu Lys Leu Glu Gly Gly Gly His Tyr Leu Cys Glu Phe Lys Ser Thr

Tyr Lys Ala Lys Lys Pro Val Met Met Pro Gly Tyr His Tyr Val Asp 180 185

Arg Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu 200

Gln Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 215

144/234

<210> 136	
<211> 663	
<212> DNA	
<213> Porites murrayensis	
<220>	
<221> CDS	
<222> (1)(663)	
<pre>&lt;400&gt; 136 atg agt gtg atc gct aca caa atg acc tac aag gtt tat atg tca ggc Met Ser Val Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Ser Gly 1 5 10 15</pre>	48
acg gtc aat gga cac tac ttt gag gtt gaa ggc gat gga aaa gga aag Thr Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys 20 25 30	96
cct tac gag ggg gag cag acg gta aag ctc act gtc acc aag ggc gga Pro Tyr Glu Gly Glu Gln Thr Val Lys Leu Thr Val Thr Lys Gly Gly 35 40 45	144
cct ctg cca ttt gct tgg gat att cta tca cca cag agt cag tac gga Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Ser Gln Tyr Gly 50 55 60	192
agc ata cca ttc acc aag tac cct gaa gac atc cct gac tat gta aag Ser Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys 65 70 75 80	240
cag tca ttc cct gag gga tat aca tgg gag agg atc atg aac ttc gaa Gln Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ile Met Asn Phe Glu 85 90 95	288
gat ggt gca gtg tgt act gtc agc aat gat tcc agc atc caa ggt aac Asp Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn 100 105 110	336
tgt ttc atc tac aat gtc aag ttc tct ggt ttg aac ttt cct ccc aat Cys Phe Ile Tyr Asn Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn 115 120 125	384
gga cct gtt atg caa aag aag aca cag ggt tgg gaa ccc aac act gag Gly Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu 130 135 140	432
cgt ctc ttt gca cga gat gga atg ctg ata gga aac aac ttt atg gct Arg Leu Phe Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala 145 150 155 160	480

## 145/234 ctg aag ttg gaa gga ggt ggt cat tat ttg tgt gaa ttc aaa tct act 528 Leu Lys Leu Glu Gly Gly Gly His Tyr Leu Cys Glu Phe Lys Ser Thr tac aag gca aag aag cct gtg atg atg cca ggg tat cac tat gtt gac 576 Tyr Lys Ala Lys Lys Pro Val Met Met Pro Gly Tyr His Tyr Val Asp 180 185 cgc aaa ttg gat gta acc aat cac aac aag gat tac act tcc gtt gag 624 Arg Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu 195 200 205 cag tgt gaa att tcc att gca cgc aaa cct gtg gtc gcc 663 Gln Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala <210> 137 <211> 221 <212> PRT <213> Porites murrayensis <400> 137 Met Ser Val Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Ser Gly 10 Thr Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys Pro Tyr Glu Gly Glu Gln Thr Val Lys Leu Thr Val Thr Lys Gly Gly Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Ser Gln Tyr Gly 55 Ser Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys 70 75 Gln Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ile Met Asn Phe Glu

Asp Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn

Cys Phe Ile Tyr Asn Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn

### 146/234

									140	1234						
Gly	Pro 130	Val	Met	Gln	Lys	Lys 135	Thr	Gln	Gly	Trp	Glu 140	Pro	Asn	Thr	Glu	
Arg 145	Leu	Phe	Ala	Arg	Asp 150	Gly	Met	Leu	Ile	Gly 155	Asn	Asn	Phe	Met	Ala 160	
Leu	Lys	Leu	Glu	Gly 165	Gly	Gly	His	Tyr	Leu 170	Суѕ	Glu	Phe	Lys	Ser 175	Thr	
Tyr	Lys	Ala	Lys 180	Lys	Pro	Val	Met	Met 185	Pro	Gly	Tyr	His	Tyr 190	Val	Asp	
Arg	Lys	Leu 195	Asp	Val	Thr	Asn	His 200	Asn	Lys	Asp	Tyr	Thr 205	Ser	Val	Glu	
Gln	Cys 210	Glu	Ile	Ser	Ile	Ala 215	Arg	Lys	Pro	Val	Val 220	Ala			ı	
<210	)> :	138														
<211	L> (	560														
<212	2> 1	ONA													•	
<213	3> 1	Pori	tes r	nurra	ayen:	sis										
<220	)>															
<221	L> (	CDS														
<222	?>	(1).	. (660	0)												
	gtg		gct Ala													. 48
			cac His 20													96
			gag Glu													144

ctg ccc ttt gct tgg gat att tta tca cct cag act cag tac gga agc Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Thr Gln Tyr Gly Ser

55

WO 02/070703	PCT/GB02/00928
11 6 02/070705	1 € 17 € 17 € 17

### 147/234

					tac Tyr 70											240
tca Ser	ttc Phe	cct Pro	gag Glu	gga Gly 85	tat Tyr	aca Thr	tgg Trp	gag Glu	agg Arg 90	atc Ile	atg Met	aag Lys	ttt Phe	gaa Glu 95	gat Asp	288
					gtc Val											336
					aag Lys											384
					aag Lys											432
					gga Gly 150											480
					ggt Gly											528
_	-	-	_		gtg Val	_							-	_	_	576
	-	_	-		aat Asn			_	_				-		_	624
-	-				gca Ala	_				_	_					660
<210	)> 1	139														

<210> 139

<211> 220

<212> PRT

<213> Porites murrayensis

<400> 139

Ser Val Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr 1  $\phantom{\bigg|}$  5  $\phantom{\bigg|}$  10  $\phantom{\bigg|}$  15

#### 148/234

Tyr Glu Gly Glu Gln Thr Val Lys Leu Thr Val Thr Lys Gly Gly Pro 35 40 45

Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Thr Gln Tyr Gly Ser 50 55 60

Ile Pro Phe Thr Lys Tyr Pro Asp Asp Ile Pro Asp Tyr Val Lys Gln 65 70 75 80

Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ile Met Lys Phe Glu Asp 85 90 95

Gly Ala Val Cys Thr Val Thr Asn Asp Ser Ser Met Gln Gly Asn Cys 100 105 110

Phe Ile Tyr Asn Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly 115 120 125

Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Gly Arg 130 135 140

Leu Tyr Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu 145 150 155 160

Lys Leu Glu Gly Gly His Tyr Thr Cys Glu Phe Lys Ser Thr Tyr
165 170 175

Lys Ala Lys Lys Pro Val Met Met Pro Gly Tyr His Tyr Val Asp Arg 180 185 190

Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu Gln
195 200 205

Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 210 215 220

<210> 140

<211> 660

<212> DNA

<213> Porites murrayensis

<220>

### 149/234

<221> CDS

<222> (1)..(660)

<400	)> ]	40															
_			-			_			-	-		_	tca Ser		_	٠	48
													gga Gly 30				96
	-			_	_	_	_			-		_	ggc Gly				144
													tac Tyr			;	192
				_			-	_			_		gta Val	-	_		240
													ttc Phe				288
													ggc Gly 110				336
			Asn										ccc Pro				384
	-	_		-	_		_			-			act Thr		_		432
													atg Met				480
		Glu	Gly	Gly		His	Tyr	Leu	Cys	Glu	Phe	Lys	tct Ser		Tyr		528
aag Lys	gca Ala	aag Lys	aag Lys 180	cct Pro	gtg Val	atg Met	atg Met	cca Pro 185	ggg Gly	tat Tyr	cac His	tat Tyr	gtt Val 190	gac Asp	cgc Arg		576
aaa Lys	ttg Leu	gat Asp 195	gta Val	acc Thr	aat Asn	cac His	aac Asn 200	aag Lys	gat Asp	tac Tyr	act Thr	tcc Ser 205	gtt Val	gag Glu	cag Gln		624

150/234

tgt gaa att tcc att gca cgc aaa cct gtg gtc gcc
Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala
210
215
220

<210> 141

<211> 220

<212> PRT

<213> Porites murrayensis

<400> 141

Ser Val Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr 1 5 10 15

Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Gln Gly Lys Pro 20 25 30

Tyr Glu Gly Glu Gln Thr Val Lys Leu Thr Val Thr Lys Gly Gly Pro 35 40 45

Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Ser Gln Tyr Gly Ser 50 60

Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys Gln 65 70 75 80

Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ile Met Asn Phe Glu Asp 85 90 95

Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys 100 105 110

Phe Ile Tyr Asn Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly 115 120 125

Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu Arg 130 135 140

Leu Phe Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu 145 150 155 160

Lys Ser Glu Gly Gly Gly His Tyr Leu Cys Glu Phe Lys Ser Thr Tyr 165 170 175

#### 151/234

Lys	Ala	Lys	Lys	Pro	Val	Met	Met	Pro	Gly	Tyr	His	Tyr	Val	Asp	Arg
			180					185					190		

Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu Gln 195 200 205

Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 210 215 220

<210> 142

<211> 660

<212> DNA

<213> Porites murrayensis

<220>

<221> CDS

<222> (1)..(660)

<400> 142					
	-	atg acc tac Met Thr Tyr			-
		gag gtt gaa Glu Val Glu 25			
5 5 555		gta aag ctc Val Lys Leu 40			
		att cta tca Ile Leu Ser 55			
	-	cct gaa gac Pro Glu Asp	_	-	
		aca tgg gag Thr Trp Glu			
		agc aat gat Ser Asn Asp 105			

WO 02/070703	PCT/GB02/0092

	***	02/0	,,,,,,	•					152	/234						C 17G D/12/002	•
ttc a Phe I	le	tac Tyr 115	aat Asn	gtc Val	aag Lys	ttc Phe	tct Ser 120	ggt Gly	ttg Leu	aac Asn	ttt Phe	cct Pro 125	ccc Pro	aat Asn	gga Gly	384	
cct g Pro V 1	tt al 30	atg Met	caa Gln	aag Lys	aag Lys	aca Thr 135	cag Gln	ggc Gly	tgg Trp	gaa Glu	ccc Pro 140	aac Asn	aca Thr	gag Glu	cgt Arg	432	
ctc t Leu P 145	tt he	gca Ala	cga Arg	gat Asp	gga Gly 150	atg Met	ctg Leu	ata Ile	gga Gly	aac Asn 155	aac Asn	ttt Phe	atg Met	gct Ala	ctg Leu 160	480	
aag t Lys L	tg eu	gaa Glu	gga Gly	ggt Gly 165	ggt Gly	cat His	tat Tyr	ttg Leu	tgt Cys 170	gaa Glu	ttc Phe	aaa Lys	tct Ser	act Thr 175	tac Tyr	528	
aag g Lys A	ca la	aag Lys	aag Lys 180	cct Pro	gtg Val	atg Met	atg Met	cca Pro 185	Gly ggg	tat Tyr	cac His	tat Tyr	gtt Val 190	gac Asp	cgc Arg	576	
aaa t Lys L	eu .	gat Asp 195	gta Val	acc Thr	aat Asn	cac His	aac Asn 200	aag Lys	gat Asp	tac Tyr	act Thr	tcc Ser 205	gtt Val	gag Glu	cag Gln	624	
tgt g Cys G 2	aa lu 10	att Ile	tcc Ser	att Ile	gca Ala	cgc Arg 215	aaa Lys	cct Pro	gtg Val	gtc Val	gcc Ala 220					660	
<210>	1	43															
<211>	2	20															
<212>	P	RT															
<213>	P	orit	es m	nurra	yens	sis											
<400>	1	43															
Ser V	al :	Ile	Ala	Thr 5	Gln	Met	Thr	Tyr	Lys 10	Val	Tyr	Met	Ser	Gly 15	Thr		
Val A	sn (	Gly	His 20	Tyr	Phe	Glu	Val	Glu 25	Gly	Asp	Gly	Lys	Gly 30	Lys	Pro		
Tyr G		Gly 35	Glu	Gln	Thr	Val	Lys 40	Leu	Thr	Val	Thr	Lys 45	Gly	Gly	Pro		
				_	_		_	_	_								

Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Ser Gln Tyr Gly Ser

Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys Gln 65 70 75 80

#### 153/234

Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ile Met Asn Phe Glu Asp 85 90 95

Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys 100 105 110

Phe Ile Tyr Asn Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly 115 120 125

Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu Arg 130 135 140

Lys Leu Glu Gly Gly Gly His Tyr Leu Cys Glu Phe Lys Ser Thr Tyr 165 170 175

Lys Ala Lys Lys Pro Val Met Met Pro Gly Tyr His Tyr Val Asp Arg 180 185 190

Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu Gln 195 200 205

Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 210 215 220

<210> 144

<211> 660

<212> DNA

<213> Pink Pocillopora

<220>

<221> CDS

<222> (1)..(660)

<400> 144

agt gtg atc gct aca caa atg acc tac aag gtt tat atg tca ggc acg Ser Val Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr 1 5 10 15

WO 02/070703	PCT/GB02/00928
--------------	----------------

#### 154/234 gtc aat gga cac tac ttt gag gtt gaa ggc gat gga aaa gga aag cct 96 Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys Pro tac gag ggc gag cag act gta aag ctc act gtc acc aag ggc gga cct 144 Tyr Glu Gly Glu Gln Thr Val Lys Leu Thr Val Thr Lys Gly Gly Pro 40 ctg ccg ttt gct tgg gat att tta tca cca cag act cag tac gga agc 192 Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Thr Gln Tyr Gly Ser ata cca ttc acc aag tac cct gaa gac att cct gac tat gta aaa cag 240 Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys Gln tca ttc cct gag gga tat aca tgg gag agg atc atg aag ttt gaa gat 288 Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ile Met Lys Phe Glu Asp 85 ggt gca gta tgt act gtc agc aat gat tcc agc atg caa ggc aac tgt 336 Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Met Gln Gly Asn Cys 105 ttc atc tac aat gtc aag ttc tct ggt ttg aac ttt cct ccc aat gga 384 Phe Ile Tyr Asn Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly 115 120 cct gtt atg cag aag aag aca cag ggc tgg gaa ccc aac act gag cgt 432 Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu Arg 135 ctt tat gca cga gat gga atg ctg ata gga aac aac ttt atg gct ctg 480 Leu Tyr Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu 150 aag ttg gaa gga ggt ggt cat tat acc tgt gaa ttc aaa tct act tac 528 Lys Leu Glu Gly Gly His Tyr Thr Cys Glu Phe Lys Ser Thr Tyr 165 aag gca aag aag cct gtg atg atg cct gga tat cac tat gtt gac cgc 576 Lys Ala Lys Lys Pro Val Met Met Pro Gly Tyr His Tyr Val Asp Arg 180 aaa ttg gat gta acc aat cac aac aag gat tac act tcc gtt gag cag 624 Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu Gln 195 200 205 tgt gaa att tcc att gca cgc aaa cct gtg gtc gcc 660 Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala

<210> 145

210

<211> 220

<212> PRT

<213> Pink Pocillopora

155/234

<400> 145

Ser Val Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr 1  $\phantom{\bigg|}$  5  $\phantom{\bigg|}$  10  $\phantom{\bigg|}$  15

Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys Pro 20 25 30

Tyr Glu Gly Glu Gln Thr Val Lys Leu Thr Val Thr Lys Gly Gly Pro 35 40

Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Thr Gln Tyr Gly Ser 50 55 60

Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys Gln 65 70 75 80

Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ile Met Lys Phe Glu Asp 85 90 95

Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Met Gln Gly Asn Cys 100 105 110

Phe Ile Tyr Asn Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly 115 120 125

Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu Arg 130 135 140

Leu Tyr Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu 145 150 160

Lys Leu Glu Gly Gly His Tyr Thr Cys Glu Phe Lys Ser Thr Tyr 165 170 175

Lys Ala Lys Lys Pro Val Met Met Pro Gly Tyr His Tyr Val Asp Arg 180 185 190

Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu Gln
195 200 205

Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 210 215 220

156/234

<210> 146 <211> 663 <212> DNA <213> Pink Pocillopora <220> <221> CDS <222> (1)..(663) <400> 146 atg agt gtg atc gct aca caa atg acc tac aag gtt tat atg tca ggc 48 Met Ser Val Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Ser Gly acg gtc aat gga cac tac ttt gag gtc gaa ggc gat gga aaa gga aag 96 Thr Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys cct tac gag ggg gag cag acg gta agg ctg gct gtc acc aag ggc gga 144 Pro Tyr Glu Gly Glu Gln Thr Val Arg Leu Ala Val Thr Lys Gly Gly 40 cct ctq cca ttt qct tgg gat att tta tca cca cag tgt cag tac gga 192 Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Cys Gln Tyr Gly 55 240 age ata cca ttc acc aag tac cct gaa gac atc cct gac tat gta aag Ser Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys 288 cag tca ttc ccg gag gga ttt aca tgg gag agg atc atg aac ttt gaa Gln Ser Phe Pro Glu Gly Phe Thr Trp Glu Arg Ile Met Asn Phe Glu gat ggt gca gtg tgt act gtc agc aat gat tcc agc atc caa ggc aac 336 Asp Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn 105 tgt ttc atc tac cat gtc aag ttc tct ggt ttg aac ttt cct ccc aat 384 Cys Phe Ile Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn 120 125 gga cet gtt atg cag aag aca cag ggc tgg gaa ccc cac tet gag 432 Gly Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro His Ser Glu 135 cgt ctc ttt gca cga gat gga atg ctg ata gga aac aac ttt atg gct 480

Arg Leu Phe Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala

# 2/00928

	wo	02/0	70703	i					157	/234						PCT/GB6	02/00
ctg a Leu L	ag	tta Leu	gaa Glu	gga Gly 165	Gly Ggc	ggt Gly	cac His	tat Tyr	ttg Leu 170	tgt Cys	gaa Glu	ttc Phe	aaa Lys	act Thr 175	act Thr	5	28
tac a Tyr L	ag ys	gca Ala	aag Lys 180	aag Lys	cct Pro	gtg Val	aag Lys	atg Met 185	cca Pro	ggg Gly	tat Tyr	cat His	tat Tyr 190	gtt Val	gac Asp	5	76
cgc a Arg L	ys	ctg Leu 195	gat Asp	gta Val	acc Thr	aat Asn	cac His 200	aac Asn	aag Lys	gat Asp	tac Tyr	act Thr 205	tcc Ser	gtt Val	gag Glu	6	24
cag t Gln C																6	63
<210>	. 1	47															
<211>	2	21								•							
<212>	P	RT															
<213>	P	ink	Poci	illop	ora												
<400>	. 1	47															
Met S 1	er	Val	Ile	Ala 5	Thr	Gln	Met	Thr	Tyr 10	Lys	Val	Tyr	Met	Ser 15	Gly		
Thr V	al .	Asn	Gly 20	His	Tyr	Phe	Glu	Val 25	Glu	Gly	Asp	Gly	Lys 30	Gly	Lys		
Pro T		Glu 35	Gly	Glu	Gln	Thr	Val 40	Arg	Leu	Ala	Val	Thr 45	Lys	Gly	Gly		
Pro L 5	eu 0	Pro	Phe	Ala	Trp	Asp 55	Ile	Leu	Ser	Pro	G <b>l</b> n 60	Cys	Gln	Tyr	Gly		
Ser I 65	le	Pro	Phe	Thr	Lys 70	Tyr	Pro	Glu	Asp	Ile 75	Pro	Asp	Туг	Val	Lys 80		
Gln S	er	Phe	Pro	Glu 85	Gly	Phe	Thr	Trp	Glu 90	Arg	Ile	Met	Asn	Phe 95	Glu		
Asp G	1у .	Ala	Val 100	Cys	Thr	Val	Ser	Asn 105	Asp	Ser	Ser	Ile	Gln 110	Gly	Asn		
												,					

Cys Phe Ile Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn 115  $\,$  120  $\,$  125

#### 158/234

Gly Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro His Ser Glu 130 135 140

Arg Leu Phe Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala 145 150 155 160

Leu Lys Leu Glu Gly Gly Gly His Tyr Leu Cys Glu Phe Lys Thr Thr 165 170 175

Tyr Lys Ala Lys Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Asp 180 185 190

Arg Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu 195 200 205

Gln Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 210 215 220

<210> 148

<211> 663

<212> DNA

<213> Pink Pocillopora

<220>

<221> CDS

<222> (1)..(663)

<400> 148

atg agt gtg atc gct aca caa atg acc tac aag gtt tat atg tca ggc

Met Ser Val Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Ser Gly

1 10 15

acg gtc aat gga cac tac ttt gag gtc gaa ggc gat gga aaa gga aag 96
Thr Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys
20 25 30

cct tac gag ggg gag cag acg gta agg ctg gct gtc acc aag ggc gga
Pro Tyr Glu Gly Glu Gln Thr Val Arg Leu Ala Val Thr Lys Gly Gly
35
40
45

cct ctg cca ttt gct tgg gat att tta tca cca cag tgt cag tac gga 192
Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Cys Gln Tyr Gly
50 55 60

WO 02/070703		PCT/GB02/00928
	159/234	

agc ata cca ttc acc Ser Ile Pro Phe Thr 65	aag tac cct o Lys Tyr Pro ( 70	gaa gac atc cct Glu Asp Ile Pro 75	gac tat gta aag Asp Tyr Val Lys 80	240
cag tca ttc ccg gag Gln Ser Phe Pro Glu 85	gga ttt aca t Gly Phe Thr 1	tgg gag agg atc Trp Glu Arg Ile 90	atg aac ttt gaa Met Asn Phe Glu 95	288
gat ggt gca gtg tgt Asp Gly Ala Val Cys 100	Thr Val Ser A	aat gat tcc agc Asn Asp Ser Ser 105	atc caa ggc aac Ile Gln Gly Asn 110	336
tgt ttc atc tac cat Cys Phe Ile Tyr His 115	gtc aag ttc t Val Lys Phe S 120	tct ggt ttg aac Ser Gly Leu Asn	ttt cct ccc aat Phe Pro Pro Asn 125	384
gga cct gtt atg cag Gly Pro Val Met Gln 130	aag aag aca d Lys Lys Thr 0 135	cag ggc tgg gaa Gln Gly Trp Glu 140	ccc cac tct gag Pro His Ser Glu	432
cgt ctc ttt gca cga Arg Leu Phe Ala Arg 145	gat gga atg o Asp Gly Met I 150	ctg ata gga aac Leu Ile Gly Asn 155	aac ttt atg gct Asn Phe Met Ala 160	480
ctg aag tta gaa gga Leu Lys Leu Glu Gly 165	ggc ggt cac t Gly Gly His T	tat ttg tgt gaa Tyr Leu Cys Glu 170	ttc aaa act act Phe Lys Thr Thr 175	528
tac aag gca aag aag Tyr Lys Ala Lys Lys 180	Pro Val Lys N			576
cgc aaa ctg gat gta Arg Lys Leu Asp Val 195				624
cag tgt gaa att tcc Gln Cys Glu Ile Ser 210				663
<210> 149				

<211> 221

<212> PRT

<213> Pink Pocillopora

<400> 149

Met Ser Val Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Ser Gly 1 5 10 15

Thr Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys 20 25 30

#### 160/234

Pro Tyr Glu Gly Glu Gln Thr Val Arg Leu Ala Val Thr Lys Gly Gly 35 40

Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Cys Gln Tyr Gly 50 55 60

Ser Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys 70 75 80

Gln Ser Phe Pro Glu Gly Phe Thr Trp Glu Arg Ile Met Asn Phe Glu 85 90 95

Asp Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn 100 105 110

Cys Phe Ile Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn 115 120 125

Gly Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro His Ser Glu 130 135 140

Arg Leu Phe Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala 145 150 155 160

Leu Lys Leu Glu Gly Gly Gly His Tyr Leu Cys Glu Phe Lys Thr Thr 165 170 175

Tyr Lys Ala Lys Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Asp 180 185 190

Arg Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu 195 200 205

Gln Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 210 215 220

<210> 150

<211> 660

<212> DNA

<213> Pink Pocillopora

<220>

# 161/234

<221> CDS

<222> (1)..(660)

<400> 150 agt gtg atc	act aca	caa ato acc	tac aag	ott tat	ato toa	aac sca	48
Ser Val Ile	Ala Thr (	Gln Met Thr	Tyr Lys	Val Tyr	Met Ser	Gly Thr	40
gtc aat gga Val Asn Gly	cac tac t His Tyr 1 20	ttt gag gtt Phe Glu Val	gaa ggc Glu Gly 25	gat gga Asp Gly	aaa gga Lys Gly 30	aag cct Lys Pro	96
tac gag ggg Tyr Glu Gly 35				Val Thr			144
ctg cca ttt Leu Pro Phe 50							192
ata cca ttc Ile Pro Phe 65	Thr Lys						240
tca ttc cct Ser Phe Pro							288
ggt gca gtg Gly Ala Val	-	-	-	_		_	336
ttc atc tac Phe Ile Tyr 115	-	-	Gly Leu	Asn Phe			384
cct gtt atg Pro Val Met 130	_			-			. 432
ctc ttt gca Leu Phe Ala 145	Arg Asp (				_		480
aag ttg gaa Lys Leu Glu				-			528
aag gca aag Lys Ala Lys							576
aaa ttg gat Lys Leu Asp 195			Lys Asp	Tyr Thr			624

660

1	67	/23	4
	υz	123	4

tgt gag att tcc att gca cgc aaa cct gtg gtc gcc
Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala
210 220

<210> 151

<211> 220

<212> PRT

<213> Pink Pocillopora

<400> 151

Ser Val Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr 1 5 10 15

Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys Pro 20 25 30

Tyr Glu Gly Glu Gln Thr Val Lys Leu Thr Val Thr Lys Gly Gly Pro 35 40 45

Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Ser Gln Tyr Gly Ser 50 55 60

Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys Gln 65 70 75 80

Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ile Met Asn Phe Glu Asp 85 90 95

Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys 100 105 110

Phe Ile Tyr Asn Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly 115 120 125

Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu Arg 130 135 140

Leu Phe Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu 145 150 155 160

Lys Leu Glu Gly Gly His Tyr Leu Cys Glu Phe Lys Ser Thr Tyr 165 170 175

#### 163/234

Lys Ala	Lys	Lys	Pro	Val	Met	Met	Pro	Gly	Tyr	His	Tyr	Val	Asp	Arg
		180					185					190		

Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu Gln 195  $\phantom{\bigg|}200\phantom{\bigg|}205\phantom{\bigg|}$ 

Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 210 215 220

<210> 152

<211> 663

<212> DNA

<213> Platygyra sp.

<220>

<221> CDS

<400> 152

<222> (1)..(663)

100

	gtg Val									48	ţ
	aat Asn									96	5
	gag Glu 35	 	 _	-	-		-	_	 	144	i
	ccg Pro									192	}
	cca Pro									240	)
	ttc Phe									288	}
	gca Ala	Cys		Ser		Asp				336	5

105

#### 164/234

									104	1234						
	ttc Phe															384
	cct Pro 130															432
	ctc Leu		-	,			-	_						_	_	480
	aag Lys															528
	aag Lys		_				_	_						-	-	576
	: aaa J Lys														gag Glu	624
-	tgt Cys 210	-				-	_				_	_				663
<21	.0>	153														
<21	.1>	221														
<21	.2>	PRT														•
<21	.3>	Platy	ygyra	a sp												
<40	0>	153														
Met 1	Ser	Val	Ile	Ala 5		Gln				_				Ser 15	<del>-</del>	
Thi	· Val	Asn	Gly 20	His	Tyr	Phe	Glu	Val 25	Glu	Gly	Asp	Gly	Lys 30	Gly	Lys	
Pro	Tyr	Glu 35	Gly	Glu	Arg	Thr	Va1 40	Lys	Leu	Thr	Val	Thr 45	Lys	Gly	Gly	
Pro	Leu 50	Pro	Phe	Ala	Trp	Asp 55	Ile	Leu	Ser	Pro	Gln 60	Cys	Gln	Tyr	Gly	
Asr 65	ılle	Pro	Phe	Thr	Lys 70	Tyr	Pro	Glu	Asp	Val 75	Pro	Asp	Tyr	Val	Lys 80	

#### 165/234

Gln Ser Phe Pro Glu Gly Phe Thr Trp Glu Arg Ile Met Asn Phe Glu 85 90 95

Asp Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn 100 105 110

Cys Phe Thr Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn 115 120 125

Gly Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro His Ser Glu 130 135 140

Arg Leu Phe Ala Arg Gly Gly Met Leu Ile Gly Asn Asn Phe Met Ala 145 150 155 160

Leu Lys Leu Glu Gly Gly Gly His Tyr Leu Cys Gly Phe Lys Thr Thr 165 170 175

Tyr Lys Ala Lys Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Asp 180 185 190

Arg Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Ile Ser Val Glu 195 200 205

Gln Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 210 215 220

<210> 154

<211> 663

<212> DNA

<213> Platygyra sp.

<220>

<221> CDS

<222> (1)..(663)

<400> 154

atg agt gtg atc gct aca caa atg acc tac aag gtt tat atg tca ggc Met Ser Val Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Ser Gly 1 5 10 15

48

#### WO 02/070703 PCT/GB02/00928 166/234

acg gtc a Thr Val A	aat gga cac Asn Gly His 20	tac ttt ga Tyr Phe G	ag gtc gaa lu Val Glu 25	ggc gat Gly Asp	gga aaa Gly Lys 30	gga aag Gly Lys	96
Pro Tyr G		g cag acg gt n Gln Thr Va 40	al Arg Leu				144
cct ctg c Pro Leu P 50	cca ttt gct Pro Phe Ala	tgg gat at Trp Asp II 55	tt ttg tca le Leu Ser	cca cag Pro Gln 60	tat cag Tyr Gln	tac gga Tyr Gly	192
		aag tac co Lys Tyr Pi 70					240
_		gga ttt ac Gly Phe Th 85			_	_	288
		acc gtc ac Thr Val Se		Ser Ser			336
tgt ttc a Cys Phe I	atc tac cat Tle Tyr His 115	gtc aag tt Val Lys Pl	tc tct ggt ne Ser Gly 120	ttg aac Leu Asn	ttt cct Phe Pro 125	ccc aat Pro Asn	384
Gly Pro V		g aag aag ad Lys Lys T) 13					432
		gat gga at Asp Gly Me 150					480
		ggt ggt ca Gly Gly Hi 165	_				528
		g cct gtg aa g Pro Val Ly )	- •				576
		acc aat ca Thr Asn Hi					624
Arg Cys G		att gca co Ile Ala An 21	•		-		663

<210> 155

<211> 80

<212> PRT

<213> Platygyra sp.

<400> 155

Met Ser Val Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Ser Gly  $1 \hspace{1cm} 5 \hspace{1cm} 10 \hspace{1cm} 15$ 

Thr Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys 20 25 30

Pro Tyr Glu Gly Glu Gln Thr Val Arg Leu Thr Val Thr Lys Gly Gly 35 40 45

Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Tyr Gln Tyr Gly 50 55 60

Ser Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys 65 70 75 80

<210> 156

<211> 140

<212> PRT

<213> Platygyra sp.

<400> 156

Ser Phe Pro Glu Gly Phe Thr Trp Asp Arg Ile Met Asn Phe Glu Asp 1 5 10 15

Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys 20 25 30

Phe Ile Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly 35 40

Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu Arg 50 55

Leu Leu Ala Arg Asp Gly Met Leu Leu Gly Asn Asn Phe Met Ala Leu 65 70 75 80

Lys Leu Glu Gly Gly His Tyr Leu Cys Glu Phe Lys Thr Thr Tyr 85 90 95

#### 168/234

Lys	Ala	Lys	Lys	Pro	Val	Lys	Met	Pro	Gly	Tyr	His	Tyr	Val	Asp	Arg
			100					105					110		

Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu Arg 115 120 125

Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 130 135 140

<210> 157

<211> 660

<212> DNA

<213> Platygyra sp.

<220>

<221> CDS

<222> (1)..(660)

100

<400> 157	
agt gtg atc gct aca caa atg acc tac aag gtt tat atg tca ggc acg Ser Val Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr 1 10 15.	
gtc aat gga cac tac ttt gag gtc gaa ggc gat gga aaa gga aag cct Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys Pro 20 25 30	
tac gag ggg gag cag acg gta aag ctc act gtc acc aag ggc gga cct Tyr Glu Gly Glu Gln Thr Val Lys Leu Thr Val Thr Lys Gly Gly Pro 35 40 45	
ctg cca ttt gct tgg gat att tta tca cca cag tgt cag tac gga aac Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Cys Gln Tyr Gly Asr 50 55 60	
ata cca ttc acc aag tac cct gaa gac gtc cct gac tat gta aag cag Ile Pro Phe Thr Lys Tyr Pro Glu Asp Val Pro Asp Tyr Val Lys Glr 65 70 75 80	
tca ttc ccg gag gga ttt aca tgg gag agg atc atg aac ttt gaa gat Ser Phe Pro Glu Gly Phe Thr Trp Glu Arg Ile Met Asn Phe Glu Asp 85 90 95	
ggt gca gtg tgt act gtc agc aat gat tcc agc atc caa ggc aac tgt	336

Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys

# PCT/GB02/00928

# WO 02/070703 169/234 ttc acc tac cat gtc aag ttc tct ggt ttg aac ttt cct ccc aat gga 384 Phe Thr Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly 115 120 cct gtg atg cag aag aca cag ggc tgg gaa ccc cac tct gag cgt 432 Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro His Ser Glu Arg 135 ctc ttt gca cgg ggt gga atg ctg ata gga aac aac ttt atg gct ctg 480 Leu Phe Ala Arg Gly Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu 150 155 aag tta gaa gga ggc ggt cac tat ttg tgt gga ttc aaa act act tac 528 Lys Leu Glu Gly Gly His Tyr Leu Cys Gly Phe Lys Thr Thr Tyr aag gca aag aag ccc gtg aag atg cca ggg tat cat tat gtt gac cgc 576 Lys Ala Lys Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Asp Arg 180 185 aaa ctg gat gta acc aat cac aac aag gat tac att tcc gtt gag cag 624 Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Ile Ser Val Glu Gln 200 tgt gaa att tcc att gca cgc aaa cct gtg gtc gcc 660 Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 215 <210> 158 <211> 220 <212> PRT <213> Platygyra sp. <400> 158 Ser Val Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr 5 10 15 Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys Pro Tyr Glu Gly Glu Gln Thr Val Lys Leu Thr Val Thr Lys Gly Gly Pro 35 40

Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Cys Gln Tyr Gly Asn

Ile Pro Phe Thr Lys Tyr Pro Glu Asp Val Pro Asp Tyr Val Lys Gln

55

50

#### 170/234 .

Ser Phe Pro Glu Gly Phe Thr Trp Glu Arg Ile Met Asn Phe Glu Asp 85 90 95

Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys 100 105 110

Phe Thr Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly 115 120 125

Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro His Ser Glu Arg 130 135 140

Leu Phe Ala Arg Gly Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu 145 150 155 160

Lys Leu Glu Gly Gly His Tyr Leu Cys Gly Phe Lys Thr Thr Tyr
165 170 175

Lys Ala Lys Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Asp Arg 180 185 190

Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Ile Ser Val Glu Gln 195 200 205

Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 210 215 220

<210> 159

<211> 660

<212> DNA

<213> Platygyra sp.

<220>

<221> CDS

<222> (1)..(660)

<400> 159

agt gtg atc gct aca caa atg acc tac aag gtt tat atg tca ggc acg Ser Val Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr 1 5 10

48

#### 171/234

gtc a Val A	at gga sn Gly	cac His 20	tac Tyr	ttt Phe	gag Glu	gtc Val	gaa Glu 25	ggc Gly	gat Asp	gga Gly	aaa Lys	gga Gly 30	aag Lys	cct Pro	96
	ag ggg lu Gly 35														144
	ca ttt ro Phe 0														192
	ca ttc ro Phe														240
	tc ccg he Pro														288
	ca gtg la Val	-		•	-		_		_					_	336
	cc tac hr Tyr 115		-	_				_							384
Pro V	tg atg al Met 30	_	_	_		_			-					-	432
	tt gca he Ala														480
_	ta gaa eu Glu						_	_							528
	ca aag la Lys														576
	tg gat eu Asp 195														624
Cys G	aa att lu Ile 10			-	-				_	_					660

<210> 160

<211> 220

<212> PRT

<213> Platygyra sp.

172/234

<400> 160

Ser Val Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr

Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys Pro 25

Tyr Glu Gly Glu Gln Thr Val Lys Leu Thr Val Thr Lys Gly Gly Pro 35

Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Cys Gln Tyr Gly Asn 55

Ile Pro Phe Thr Lys Tyr Pro Glu Asp Val Pro Asp Tyr Val Lys Gln 70

Ser Phe Pro Glu Gly Phe Thr Trp Glu Gly Ile Met Asn Phe Glu Asp

Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys 105

Phe Thr Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly 115

Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro His Ser Glu Arg 135

Leu Phe Ala Arg Gly Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu 150 155

Lys Leu Glu Gly Gly His Tyr Leu Cys Gly Phe Lys Thr Thr Tyr

Lys Ala Lys Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Asp Arg

Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Ile Ser Val Glu Gln

Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 215

# WO 02/070703 PCT/GB02/00928 173/234

<210> 161	
<211> 660	
<212> DNA	
<213> Pavona decussata	
<220>	
<221> CDS	
<222> (1)(660)	
<400> 161 agt gtg atc gct aca caa atg acc tac aag gtt tat atg tca ggc acg	48
Ser Val Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr  1 5 10 15	40
gtc aat gga cac tac ttt gag gtt gaa ggc gat gga aaa gga gag cct	96
Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Glu Pro	30
tac gag ggg gag cag acg gta agg ctc act gtc aca aag ggc gga cct	144
Tyr Glu Gly Glu Gln Thr Val Arg Leu Thr Val Thr Lys Gly Gly Pro 35 40	
ctg cca ttt gct tgg gat att tta tca cca cag tat cag tac gga agc	192
Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Tyr Gln Tyr Gly Ser 50 55 60	
ata cca ttc acc aag tac cct gaa gac atc cct gac tat gta aag cag	240
Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys Gln 65 70 75 80	
tca ttc ccg gaa gga tat aca tgg gag agg atc atg aac ttt gaa gat	288
Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ile Met Asn Phe Glu Asp 85 90 95	
ggt gct gtg tgt act gtc agc aat gat tcc agc atc caa ggc aac tgt Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys	336
100 105 110	
ttc atc tac cat gtc aag ttc tct ggt ttg aac ttt cct ccc aat gga Phe Ile Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly	384
115 120 125	
cct gtg acg cag aag aca cag ggc tgg gaa ccc aac act gag cgt Pro Val Thr Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu Arg	432
130 135 140	
ctc ttt gca cga gat gga atg ctg ata gga aac aac ttt atg gct ctg Leu Phe Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu	480
145 150 155 160	

WO 02/070703	PCT/GB02/00928

			174/	234			
aag tta gaa q Lys Leu Glu (			-	_	•		528
aag gca aag a Lys Ala Lys I		Lys Met					576
aaa ctg gat o Lys Leu Asp V 195							624
tgt gaa att t Cys Glu Ile S 210	Ser Ile Ala	-					, 660
<210> 162							
<211> 220							
<212> PRT							
<213> Pavona	a decussata						
<400> 162							
Ser Val Ile <i>I</i> 1	Ala Thr Gln 5	Met Thr	Tyr Lys 10	Val Tyr	Met Ser	Gly Thr 15	
Val Asn Gly I	His Tyr Phe 20		Glu Gly 25	Asp Gly	Lys Gly 30	Glu Pro	
Tyr Glu Gly (	Glu Gln Thr	Val Arg : 40	Leu Thr	Val Thr	Lys Gly 45	Gly Pro	
Leu Pro Phe 1	Ala Trp Asp	Ile Leu 55	Ser Pro	Gln Tyr 60	Gln Tyr	Gly Ser	
Ile Pro Phe 5	Thr Lys Tyr 70	Pro Glu .	Asp Ile	Pro Asp 75	Tyr Val	Lys Gln 80	
Ser Phe Pro (	Glu Gly Tyr 85	Thr Trp	Glu Arg 90	Ile Met	Asn Phe	Glu Asp 95	
Gly Ala Val (	Cys Thr Val		Asp Ser 105	Ser Ile	Gln Gly 110	Asn Cys	
Phe Ile Tyr 1	His Val Lys	Phe Ser	Gly Leu	Asn Phe	Pro Pro 125	Asn Gly	

W	O 02/0	70703	5					175	/234						PCT/GB02/0
Pro Val		Gln	Lys	Lys	Thr 135	Gln	Gly	Trp	Glu	Pro 140	Asn	Thr	Glu	Arg	
Leu Phe 145	Ala	Arg	Asp	Gly 150	Met	Leu	Ile	Gly	Asn 155	Asn	Phe	Met	Ala	Leu 160	
Lys Leu	Glu	Gly	Gly 165	Gly	His	Tyr	Leu	Cys 170	Glu	Phe	Lys	Ser	Thr 175	Tyr	
Lys Ala	Lys	Lys 180	Thr	Val	Lys	Met	Pro 185	Gly	Tyr	His	Tyr	Val 190	Asp	Arg	
Lys Leu	Asp 195	Val	Thr	Asn	His	Asn 200	Lys	Asp	Tyr	Thr	Ser 205	Val	Glu	Gln	
Cys Glu 210		Ser	Ile	Ala	Arg 215	Lys	Pro	Val	Val	Ala 220					
<210>	163														
<211>	663														
<212>	DNA		•												
<213>	Pavo	na de	ecuss	sata											
<220>															
<221>	CDS														
<222>	(1).	. (66	3)												
<400> atg agt Met Ser 1	gtg														48
acg gtc Thr Val															96
cct tac Pro Tyr															144

cct ctg cca ttt gct tgg gat att tta tca cca cag tat cag tac gga  $Pro\cdot Leu$  Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Tyr Gln Tyr Gly 50

192

WO 02/070703	PCT/GB02/00928

### 176/234 age ata cea the ace aag tae cet gaa gae ate eet gae tat gta tag 240 Ser Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val cag tea tte eeg gaa gga tat aca tgg gag agg ate atg aac ttt gaa 288 Gln Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ile Met Asn Phe Glu gat ggt gct gtg tgt act gtc agc aat gat tcc agc atc caa ggc aac 336 Asp Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn 100 105 tgt ttc atc tac cat gtc aag ttt tct ggt ttg aac ttt cct ccc aat 384 Cys Phe Ile Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn gga cct gtg atg cag aag aag aca cag ggc tgg gaa ccc aac act gag 432 Gly Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu 135 cgt ctc ttt gca cga gat gga ttg ctg ata gga aac aac ttt atg gct 480 Arg Leu Phe Ala Arg Asp Gly Leu Leu Ile Gly Asn Asn Phe Met Ala 150 ctg aag tta gaa gaa ggc ggt cac tat ttg tgt gaa ttc aaa tcg act 528 Leu Lys Leu Glu Glu Gly Gly His Tyr Leu Cys Glu Phe Lys Ser Thr 165 170 tac aag gca aag aag act gcg aag atg cca ggg tat cac tat gtt gac 576 Tyr Lys Ala Lys Lys Thr Ala Lys Met Pro Gly Tyr His Tyr Val Asp 185 cgc aaa ctg gat gta acc aat cac aac aag gat tac act tcc gtt gag 624 Arg Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu 195 200 cag tgt gaa att tcc att gca cgc aaa cct gtg gtc gcc 663 Gln Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala <210> 164

<211>

<212> PRT

<213> Pavona decussata

<400> 164

Met Ser Val Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Ser Gly

Thr Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys

#### 177/234

Pro Tyr Glu Glu Gln Thr Val Arg Leu Thr Val Thr Lys Gly Gly 35 40 45

Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Tyr Gln Tyr Gly 50 55 60

Ser Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val 65 70 75

<210> 165

<211> 141

<212> PRT

<213> Pavona decussata

<400> 165

Gln Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ile Met Asn Phe Glu 1 5 10 15

Asp Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn 20 25 30

Cys Phe Ile Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn 35 40 45

Gly Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu 50 60

Arg Leu Phe Ala Arg Asp Gly Leu Leu Ile Gly Asn Asn Phe Met Ala 65 70 75 80

Leu Lys Leu Glu Glu Gly Gly His Tyr Leu Cys Glu Phe Lys Ser Thr 85 90 95

Tyr Lys Ala Lys Lys Thr Ala Lys Met Pro Gly Tyr His Tyr Val Asp 100 105 110

Arg Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu 115 120 125

Gln Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 130 135

# 178/234

<21	0>	166														•	
<21	1>	663															
<21	2>	DNA															
<21	3>	Pavo	na d	ecus	sata												
<22	0>														•		
<22	1>	CDS															
<22	222> (1)(663)																
<400 atg		166 gtg	atc	gct	aca	caa	gtg	acc	tac	aag	att	tat	ato	tca	aac	4	8
Met 1	Ser	Val	Ile	Āla 5	Thr	Gln	Vaĺ	Thr	Tyr 10	Lys	Val	Tyr	Met	Ser 15	Gly	3	
acg Thr	gtc Val	aat Asn	gga Gly	cac His	tac Tyr	ttt Phe	gag Glu	gtt Val	gaa Glu	ggc Glv	gat Asp	gga Glv	aaa Lvs	gga Glv	aag Lvs	9	6
		-	20		_			25		,	-	_	30	,	-1~		
cct Pro	tac Tyr	gag Glu	ggg Gly	gag Glu	caa Gln	acg Thr	gta Val	agg Arg	ctc Leu	act Thr	gtc Val	aca Thr	aag Lys	ggc Gly	gga Gly	14	4
		35					40					45					
Pro	Leu 50	cca Pro	Phe	Ala	Trp	gat Asp 55	Ile	tta Leu	tca Ser	cca Pro	Cag Gln 60	tat Tyr	cag Gln	tac Tyr	gga Gly	19	2
agc		cca	ttc	acc	aaq		cct	αaa	gac	atc		gac	tat	ata	aad	24	Λ
Ser 65	Ile	Pro	Phe	Thr	Lys 70	Tyr	Pro	Glu	Asp	Ile 75	Pro	Asp	Tyr	Val	Lys 80	24	•
cag	tca	ttc	ccg	gaa	gga	tat	aca	tgg	gag	agg	atc	atg	aac	ttt	gaa	288	8
Gln	Ser	Phe	Pro	Glu 85	Gly	Tyr	Thr	Trp	Glu 90	Arg	Ile	Met	Asn	Phe 95	Glu		
gat	ggt Glv	gct Ala	gtg Val	tgt Cvs	act	gtc	agc	aat	gat	tcc	agc	atc	caa	ggc	aac	33	6
пор	OL y	1114	100	Cys	1111	Vai	Ser	105	ASP	261	ser	ше	110	стА	Asn		
tgt Cys	ttc Phe	atc Ile	tac Tyr	cat His	gtc Val	aag Lvs	ttt Phe	tct Ser	ggt Glv	ttg Leu	aac Asn	ttt Phe	cct	ccc	aat Asn	384	4
_		115	-			-	120		3			125		110			
gga Gly	cct Pro	gtg Val	atg Met	cag Gln	aag Lys	aag Lys	aca Thr	cag Gln	ggc Gly	tgg Trp	gaa Glu	ccc Pro	aac Asn	act Thr	gag Glu	432	2
	130					135					140						
Arg	ctc Leu	ttt Phe	gca Ala	cga Arg	Asp	gga Gly	atg Met	ctg Leu	ata Ile	Gly	aac Asn	aac Asn	ttt Phe	atg Met	gct Ala	480	)
145					150					155					160		

WO 02/070703	PCT/GB02/00928

	wo	02/0	70703	;					,	PCT/G	B02/00						
ctg Leu	aag Lys	tta Leu	gaa Glu	gga Gly 165	ggc Gly	ggt Gly	cac His	tat Tyr	179/ ttg Leu 170	tgt	gaa Glu	ttc Phe	aaa Lys	tcg Ser 175	act Thr		528
tac Tyr	aag Lys	gca Ala	aag Lys 180	aag Lys	act Thr	gtg Val	aag Lys	atg Met 185	cca Pro	Gly aaa	tat Tyr	cac His	tat Tyr 190	gtt Val	gac Asp		576
					acc Thr												624
cag Gln	tgt Cys 210	gaa Glu	att Ile	tcc Ser	att Ile	gca Ala 215	cgc Arg	aaa Lys	cct Pro	gtg Val	gtc Val 220	gcc Ala					663
<210	O> :	L 67															
<21	1> 2	221															
<212	2> 1	PRT															
<213	3> I	Pavor	na de	ecus	sata												
<400	0> :	167															
Met 1	Ser	Val	Ile	Ala 5	Thr	Gln	Val	Thr	Tyr 10	Lys	Val	Tyr	Met	Ser 15	Gly		
Thr	Val	Asn	Gly 20	His	Tyr	Phe	Glu	Val 25	Glu	Gly	Asp	Gly	Lys 30	Gly	Lys		
Pro	Tyr	Glu 35	Gly	Glu	Gln	Thr	Val 40	Arg	Leu	Thr	Val	Thr 45	Lys	Gly	Gly		
Pro	Leu 50	Pro	Phe	Ala	Trp	Asp 55	Ile	Leu	Ser	Pro	Gln 60	Tyr	Gln	Tyr	Gly		

Ser Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys 

Gln Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ile Met Asn Phe Glu 

Asp Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn 

Cys Phe Ile Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn 

# 180/234

Gly Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu 130 135 140	
Arg Leu Phe Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala 145 150 155 160	
Leu Lys Leu Glu Gly Gly His Tyr Leu Cys Glu Phe Lys Ser Thr 165 170 175	
Tyr Lys Ala Lys Lys Thr Val Lys Met Pro Gly Tyr His Tyr Val Asp 180 185 190	
Arg Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu 195 200 205	
Gln Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 210 215 220	
<210> 168	
<211> 660	
<212> DNA	
<213> Pavona decussata	
<220>	
<221> CDS	
<222> (1)(660)	
<400> 168	
agt gtg atc gct aca caa atg acc tac aag gtt tat atg tca ggc acg Ser Val Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr 1 5 10 15	48
gtc aat gga cac tac ttt gag gtt gaa ggc gat gga aaa gga aag cct Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys Pro 20 25 30	96
tac gag ggg gag cag acg gta agg ctc act gtc aca aag ggc gga cct Tyr Glu Gly Glu Gln Thr Val Arg Leu Thr Val Thr Lys Gly Gly Pro 35 40 45	44
ctg cca ttt gct tgg gat att tta tca cca cag tat cag tac gga agc Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Tyr Gln Tyr Gly Ser 50 55 60	.92

WO 02/070703	PCT/GB02/00928

									181	/234						
ata Ile 65	cca Pro	ttc Phe	acc Thr	aag Lys	tac Tyr 70	cct Pro	gaa Glu	gac Asp	atc Ile	cct Pro 75	gac Asp	tat Tyr	gta Val	aag Lys	cag Gln 80	240
tca Ser	ttc Phe	ccg Pro	gaa Glu	gga Gly 85	tat Tyr	aca Thr	tgg Trp	gag Glu	ggg Gly 90	atc Ile	atg Met	aac Asn	ttt Phe	gaa Glu 95	gat Asp	288
ggt Gly	gct Ala	gtg Val	tgt Cys 100	act Thr	gtc Val	agc Ser	aat Asn	gat Asp 105	tcc Ser	agc Ser	atc Ile	caa Gln	ggc Gly 110	aac Asn	tgt Cys	336
ttc Phe	atc Ile	tac Tyr 115	cat His	gtc Val	aag Lys	ttc Phe	tct Ser 120	ggt Gly	ttg Leu	aac Asn	ttt Phe	cct Pro 125	ccc Pro	aat Asn	gga Gly	384
cct Pro	gtg Val 130	atg Met	cag Gln	aag Lys	aag Lys	aca Thr 135	cag Gln	ggc	tgg Trp	gaa Glu	ccc Pro 140	aac Asn	act Thr	gag Glu	cgt Arg	432
		gca Ala														480
		gaa Glu														528
aag Lys	gca Ala	aag Lys	aag Lys 180	act Thr	gtg Val	aag Lys	atg Met	cca Pro 185	ggg	tat Tyr	cac His	tat Tyr	gtt Val 190	gac Asp	cgc Arg	576
aaa Lys	ctg Leu	gtt Val 195	gta Val	acc Thr	aat Asn	cac His	aac Asn 200	aag Lys	gat Asp	tac Tyr	act Thr	tcc Ser 205	gtt Val	gag Glu	cag Gln	624
tgt Cys	gaa Glu 210	att Ile	tcc Ser	att Ile	gca Ala	cgc Arg 215	aaa Lys	cct Pro	gtg Val	gtc Val	gcc Ala 220					660
<210	)> :	169														
<211	L> :	220														

<212> PRT

<213> Pavona decussata

<400> 169

Ser Val Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr 1  $\phantom{\bigg|}$  5  $\phantom{\bigg|}$  10  $\phantom{\bigg|}$  15

Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys Pro

#### 182/234

Tyr Glu Gly Glu Gln Thr Val Arg Leu Thr Val Thr Lys Gly Gly Pro 35 40 45

Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Tyr Gln Tyr Gly Ser 50 55 60

Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys Gln 65 70 75 80

Ser Phe Pro Glu Gly Tyr Thr Trp Glu Gly Ile Met Asn Phe Glu Asp 85 90 95

Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys 100 105 110

Phe Ile Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly 115 120 125

Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu Arg 130 135 140

Leu Phe Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu 145 150 155 160

Lys Leu Glu Gly Gly His Tyr Leu Cys Glu Phe Lys Ser Thr Tyr 165 170 175

Lys Ala Lys Lys Thr Val Lys Met Pro Gly Tyr His Tyr Val Asp Arg

Lys Leu Val Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu Gln 195 200 205

Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 210 215 220

<210> 170

<211> 663

<212> DNA

<213> Montipora sp.

<220>

#### 183/234

<221> CDS

<222> (1)..(663)

<40 atg Met	agt	170 gtg Val	atc Ile	gct Ala	aca Thr	caa Gln	atg Met	acc Thr	tac Tvr	aag Lvs	gtt Val	tat Tvr	atg Met	tca Ser	ggc Gl v	4	18
1				5					10	-10		- ] -		15	019		
acg Thr	gtc Val	aat Asn	gga Gly 20	cac His	tac Tyr	ttt Phe	gag Glu	gtt Val 25	gaa Glu	ggc Gly	gat Asp	gga Gly	aaa Lys 30	gga Gly	aag Lys	S	96
cct Pro	tac Tyr	gaa Glu 35	Gly ggg	gag Glu	cag Gln	acg Thr	gta Val 40	agg Arg	ctc Leu	act Thr	gtc Val	aca Thr 45	aag Lys	ggc Gly	gga Gly	. 14	14
cct Pro	ctg Leu 50	cca Pro	ttt Phe	gct Ala	tgg Trp	gat Asp 55	att Ile	tta Leu	tca Ser	cca Pro	cag Gln 60	tat Tyr	cag Gln	tac Tyr	gga Gly	19	92
agc Ser 65	ata Ile	cca Pro	ttc Phe	acc Thr	aag Lys 70	tac Tyr	cct Pro	gaa Glu	gac Asp	atc Ile 75	cct Pro	gac Asp	tat Tyr	gta Val	aag Lys 80	24	10
cag Gln	tca Ser	ttc Phe	ccg Pro	gaa Glu 85	gga Gly	tat Tyr	aca Thr	tgg Trp	gag Glu 90	agg Arg	atc Ile	atg Met	aac Asn	ttt Phe 95	gaa Glu	28	38
gat Asp	ggt Gly	gca Ala	gtg Val 100	tgt Cys	gct Ala	gtc Val	agc Ser	aat Asn 105	gat Asp	tcc Ser	agc Ser	atc Ile	caa Gln 110	ggc Gly	aac Asn	33	36
tgt Cys	ttc Phe	atc Ile 115	tac Tyr	cat His	gtc Val	aag Lys	ttc Phe 120	tct Ser	ggt Gly	ttg Leu	aac Asn	ttt Phe 125	cct Pro	ccc Pro	aat Asn	38	3 4
		gtg Val														43	32
		ttt Phe														48	80
		tta Leu														52	8
		gca Ala														57	6
		ctg Leu 195														62	24

184/234

cag tgt gaa att tcc att gca cgc aaa cct gtg gtc gcc 663
Gln Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala
210 215 220

<210> 171

<211> 221

<212> PRT

<213> Montipora sp.

<400> 171

Met Ser Val Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Ser Gly
1 5 10 15

Thr Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys 20 25 30

Pro Tyr Glu Gly Glu Gln Thr Val Arg Leu Thr Val Thr Lys Gly Gly 35 40 45

Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Tyr Gln Tyr Gly 50 55 60

Ser Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys 65 70 75 80

Gln Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ile Met Asn Phe Glu 85 90 95

Asp Gly Ala Val Cys Ala Val Ser Asn Asp Ser Ser Ile Gln Gly Asn 100 105 110

Cys Phe Ile Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn 115 120 125

Gly Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu 130 135 140

Arg Leu Phe Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala 145 150 155 160

Leu Lys Leu Glu Gly Gly Gly His Tyr Leu Cys Glu Phe Lys Ser Thr 165 170 175

#### 185/234

Tyr Lys Ala	Lys Lys Pro	Val Lys	Met Pro	Gly Tyr F	His Tyr Val Asp
	180		185		190

Gln Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 210 215 220

<210> 172

<211> 663

<212> DNA

<213> Montipora sp.

<220>

<221> CDS

<222> (1)..(663)

<4	00> :	172										
	g agt Ser		-			_		_	-	-		48
	g gtc r Val											96
	tac Tyr											144
	ctg Leu 50											192
	c ata c Ile											240
	g tca n Ser											288
-	t ggt o Gly	-		_	•	-	_		-			336

	WO	0210	/0/03	•					186	/234					,	FC 17G D02/009
tgt t Cys I																384
gga c Gly E	ect Pro 130	gtg Val	atg Met	caa Gln	aaa Lys	aag Lys 135	aca Thr	caa Gln	ggc Gly	tgg Trp	gaa Glu 140	ccc Pro	aac Asn	act Thr	gag Glu	432
cgt o Arg I 145	ctc Leu	ttt Phe	gca Ala	cga Arg	gat Asp 150	gga Gly	atg Met	ctg Leu	ata Ile	gga Gly 155	aac Asn	aac Asn	ttt Phe	atg Met	gct Ala 160	480
ctg a Leu I	aag Lys	tta Leu	gaa Glu	gga Gly 165	ggc Gly	ggt Gly	cac His	tat Tyr	ttg Leu 170	tgt Cys	gaa Glu	ttc Phe	aaa Lys	tct Ser 175	act Thr	528
tac a Tyr I																576
cgc a Arg I	aaa Lys	ctg Leu 195	gat Asp	gta Val	acc Thr	aat Asn	cac His 200	aac Asn	aag Lys	gat Asp	tac Tyr	act Thr 205	tcc Ser	gtt Val	GJ À GGG	624
cag t Gln C														•		663
<210>	> 1	73														
<211>	> 2	21														
<212>	> P	RT														
<213>	> M	lonti	pora	a sp	•											
<400>	> 1	.73														
Met S	Ser	Val	Ser	Ala 5	Thr	Gln	Met	Thr	Tyr 10	Lys	Val	Tyr	Met	Ser 15	Gly	
Thr V	/al	Asn	Gly 20	His	Tyr	Phe	Glu	Val 25	Glu	Gly	Asp	Gly	Lys 30	Gly	Lys	
Pro T	Tyr	Glu 35	Gly	Glu	Gln	Thr	Val 40	Arg	Leu	Thr	Val	Thr 45	Lys	Gly	Gly	
Pro L	Leu 50	Pro	Phe	Ala	Trp	Asp	Ile	Leu	Ser	Pro	Gln	Tyr	Gl'n	Tyr	Gly	

Ser Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Gly Tyr Val Lys 65 70 75 80

#### 187/234

Gln Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ile Met Asn Phe Glu 85 90 95

Asp Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn 100 105 110

Cys Phe Ile Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn 115 120 125

Gly Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu 130 135 140

Arg Leu Phe Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala 145 150 155 160

Leu Lys Leu Glu Gly Gly Gly His Tyr Leu Cys Glu Phe Lys Ser Thr 165 170 175

Tyr Lys Ala Lys Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Asp 180 185 190

Arg Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Gly 195 200 205

Gln Cys Glu Ile Ser Ile Ala Pro Lys Pro Val Val Ala 210 215 220

<210> 174

<211> 660

<212> DNA

<213> Montipora sp.

<220>

<221> CDS

<222> (1)..(660)

<400> 174

agt gtg atc gct aca caa atg acc tac aag gtt tat atg tca ggc acg
Ser Val Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr
1 5 10

WO 02/070703		PCT/GB02/00928
	188/234	

gtc Val	aat Asn	gga Gly	cac His 20	tac Tyr	ttt Phe	gag Glu	gtc Val	gaa Glu 25	ggc Gly	gat Asp	gga Gly	aaa Lys	gga Gly 30	aag Lys	cct Pro	96
tac Tyr	gag Glu	ggg Gly 35	gag Glu	cag Gln	acg Thr	gta Val	aag Lys 40	ctc Leu	act Thr	gtc Val	acc Thr	aag Lys 45	ggc Gly	gga Gly	cct Pro	144
ctg Leu	cca Pro 50	ttt Phe	gct Ala	tgg Trp	gat Asp	att Ile 55	tta Leu	tca Ser	cca Pro	cag Gln	tgt Cys 60	cag Gln	tac Tyr	gga Gly	agc Ser	192
ata I <b>l</b> e 65	cca Pro	ttc Phe	acc Thr	aag Lys	tac Tyr 70	cct Pro	gaa Glu	gac Asp	atc Ile	cct Pro 75	gac Asp	tat Tyr	gta Val	aag Lys	cag Gln 80	240
		ccg Pro														288
		gtg Val														336
		tac Tyr 115														384
		atg Met														432
		gca Ala														480
		gaa Glu														528
		aag Lys														576
		gat Asp 195														624
		att Ile														660

<210> 175

<211> 220

<212> PRT

<213> Montipora sp.

<400> 175

Ser Val Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr 10

Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys Pro

Tyr Glu Gly Glu Gln Thr Val Lys Leu Thr Val Thr Lys Gly Gly Pro 40

Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Cys Gln Tyr Gly Ser 55 . 60

Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys Gln . 70 75

Ser Phe Pro Glu Gly Phe Thr Trp Glu Arg Ile Met Asn Phe Glu Asp

Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys

Phe Thr Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly 115 120

Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro His Ser Glu Arg

Leu Phe Ala Arg Gly Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu

Lys Leu Glu Gly Gly His Tyr Leu Cys Glu Phe Lys Thr Thr Tyr 165 170

Lys Ala Lys Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Asp Arg 185

Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu Gln

Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala

190/234

<210>	> 1	.76														
<211>	> 6	60														
<212>	> [	ANG														
<213>	> N	iont	ipora	a sp	-											
<220>	>															
<221>	<221> CDS															
<222> (1)(660)																
<400> agt g		.76 atc	att	aca	caa	ato	acc	tac	aaa	att	tat	atq	tca	aac	aca	48
Ser V	Jal	Ile	Val	Thr 5	Gln	Met	Thr	Tyr	Lys 10	Val	Tyr	Met	Ser	Gly 15	Thr	
gtc a	aat	gga	cac	tac	ttt	gag	gtt	gaa	ggc	gat	gga	aaa	gga	aag	cct	96
Val A	Asn	Gly	His 20	Tyr	Phe	Glu	Val	Glu 25	Gly	Asp	Gly	Lys	Gly 30	Lys	Pro	
tac g	gaa	ggg	gag	cag	acg	gta	agg	ctc	act	gtc	aca	aag	ggc	gga	ccc	144
Tyr G	31u	35 35	GIu	GIn	Thr	Val	Arg 40	Leu	Thr	Val	Thr	Lys 45	Gly	Gly	Pro	•
ctg c	cca	ttt	gct	tgg	gat	att	tta	tca	cca	cag	tat	cag	tac	gga	agc	192
	50	rne	лта	пр	лэр	55	neu	261	FIO	GIII	60	GIII	171	сту	ser	
ata c																240
65					70					75		-3-		-1-	80	
tca t Ser F																288
				85			-		90					95	•	
ggt g Gly A								Asp								336
			100					105					110			
ttc a Phe I	itc [le	Tyr	His	gtc Val	aag Lys	ttc Phe	Ser	ggt Gly	ttg Leu	aac Asn	ttt Phe	Pro	ccc Pro	aat Asn	gga Gly	384
		115					120					125				420
cct g Pro V																432
ctc t		gca	caa	gat	aas		cta	ata	gga	aac		+++	ata	act	cta	480
Leu P																900
										_ , ,						

#### PCT/GB02/00928 WO 02/070703

191/234												
aag tta gaa Lys Leu Glu											528	
aag gca aag Lys Ala Lys				Gly							576	
aaa ctg gat Lys Leu Asp 195											624	
tgt gaa att Cys Glu Ile 210		_			_	-					660	
<210> 177												
<211> 220												
<212> PRT												
<213> Mont	ipora sp	•										
<400> 177		·										
Ser Val Ile 1	Val Thr 5	Gln Met	Thr Ty	r Lys 10	Val	Tyr	Met	Ser	Gly 15	Thr		
Val Asn Gly	His Tyr 20	Phe Glu	Val Gl	u Gly	Asp	Gly	Lys	Gly 30	Lys	Pro		
Tyr Glu Gly 35	Glu Gln	Thr Val	Arg Le	u Thr	Val	Thr	Lys 45	Gly	Gly	Pro		
Leu Pro Phe 50	Ala Trp	Asp Ile 55	Leu Se	r Pro	Gln	Tyr 60	Gln	Tyr	Gly	Ser		
Ile Pro Phe 65	Thr Lys	Tyr Pro 70	Glu As	o Ile	Pro 75	Asp	Tyr	Val	Lys	Gln 80		
Ser Phe Pro	Glu Gly 85	Tyr Thr	Trp Gl	u Arg 90	Ile	Met	Asn	Phe	Glu 95	Asp		
Gly Ala Val	Cys Thr 100	Val Ser	Asn As	_	Ser	Ile	Gln	Gly 110	Asn	Cys		
Phe Ile Tyr 115	His Val	Lys Phe	Ser Gl 120	y Leu	Asn	Phe	Pro 125	Pro	Asn	Gly		

#### 192/234

Pro	Val	Met	Gln	Lys	Lys	Thr	Gln	Gly	Trp	Glu	Pro	Asn	Thr	Glu	Ara
	130					135					140				,

Leu Phe Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu 145 150 155 160

Lys Leu Glu Gly Gly His Tyr Leu Cys Glu Phe Lys Ser Thr Tyr 165 170 175

Lys Ala Lys Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Asp Arg 180 185 190

Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu Gln 195 200 205

Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 210 215 220

<210> 178

<211> 701

<212> DNA

<213> Acanthastria sp.

<400> 178

teegttateg etaaacagat gacegettea aegttaagtt gacaacagga agcaegaegg 60 agactgcagt cccgtacgcg cgaacgggat acctgggatt tatcaagaga acagatttca 120 cgcagacaga tggagcccgg catgacgcgt tatttgtggt tggccctctt gaagaaacca 180 tgatattgcg tggtatgagg tatcacccgg tagatatcga gaacacagtg acgagatgtc 240 atcgatcaat ctgtgaaagt gcggtcttca cgatgacaaa cctacttgtg gtagcagtgg 300 agettgatge agatgaacge gaggeacttg aegtggttee getggtgaeg acateegtae 360 tgaatgaaca gcaacttgtc gtaggggtgg tggtagtggt tgaccctggc gtagtcccga 420 tcaattctcg cggagagaaa caacggatgc atctgaggga cgggttcctg ggggaccagt 480 tggatcctat ctacgtggcg tataatatgt agacacctca ctqcttaqtt tcgtaattga 540 attgtgtcgt agttttttta aatgacaatt aatagacaag tttgaaattg actgtagcgc 600 taggtttagg tataaactag cgtttggtaa ggcaattatg acaggaacta ctgtcacgcg 660 tgacgcgaga ccgtcacttt acacgcaaac ctgtggtcgc c 701

# WO 02/070703 PCT/GB02/00928 193/234

<210> 179

<211> 701

<212> DNA

<213> Green Pocillopora

<400> 179 teegttateg etaaacagat gacegettea aegttaagtt gacaacagga ageacgaegg 60 agactgcagt cccgtacgcg cgaacgggat acctgggatt tatcaagaga acagatttca 120 cgcagacaga tggagcccgg catgacgcgt tatttgtggt tggccctctt gaaqaaacca 180 tgatattgcg tggtatgagg tatcacccgg tagatatcga gaacacagtg acgagatgtc 240 atcgatcaat ctgtgaaagt gcggtcttca cgatgacaaa cctacttgtg gtagcagtgg 300 agettgatge agatgaacge gaggeacttg acgtggttee getggtgacg acateegtae 360 tgaatgaaca gcaacttgtc gtaggggtgg tggtagtggt tgaccctggt gtagtcccga 420 tcaattctcg cggagagaaa caacggatgc atctgaggga cgggttcctg ggggaccagt 480 tggatcctat ctacgtggcg tataatatgt agacacctca ctgcttaatt ttcgtaattg 540 aattgtgtcg tagtttttt aaatgacaac taatagacag tttgaaattg actgtagcgc 600 taggtttagg tataaactag cgtttggtaa ggcaattatg acaggaatta ctgtcacgcg 660 tgacgcgaga ccgtcacttt acacgcaaac ctgtggtcgc c 701

<210> 180

<211> 701

<212> DNA

<213> Green Pocillopora

<220>

<221> misc\_feature

<222> (634)...(634)

<223> n = any nucleotide

<220>

#### WO 02/070703 PCT/GB02/00928 194/234

<221> misc\_feature
<222> (640)..(640)
<223> n = any nucleotide

<400> 180 teegttateg etaaacagat gacegettea aegttaagtt gacaacagga agcaeggegg 60 agactgcagt cccgtacgcg cgaacgggat acctgggatt tatcaagaga acagatttca 120 cgcagacaga tggagccgg catgacgcgt tatttgtggt tggccctctt gaaqaaacca 180 tgatattgcg tggtatgagg tatcacccgg tagatatcga gaacacagtg acqagatqtc 240 atcgatcaat ctgtgaaagt gcggtcttca cgatgacaaa cctacttgtg gtagcagtgg 300 agettgatge agatgaacge gaggeacttg acgtggttee getggtgaeg acateegtae 360 tgaatgaaca gcaacttgtc gtaggggtgg tggtagtggt tgaccctggc gtagtcccga 420 tcaattctcg cggagagaaa caacggatgc atctgaggga cgggttcctg ggggaccagt 480 tggatcctat ctacgtggcg tataatatgt agacacctca ctgcttagtt tcqtaattqa 540 attgtgtcgt agtttttta aatgacaatt aatagacaag tttgaaattg actgtagcqc 600 taggtttagg tataaactag cgtttggtaa ggcnattatn acaggaacta ctgtcacqcq 660 tgacgcgaga ccgtcacttt acacgcaaac ctgtggtcgc c 701

<210> 181

701

<211>

<212> DNA

<213> Green Pocillopora

<400> 181
tccgttatcg ctaaacagat gaccgcttca ccgttaagtt gacaacagga agcacgacgg 60
agactgcagt cccgtacgcg cgaacgggat acctgggatt tatcaagaga acagatttca 120
cgcagacagg tggagcccgg catgacgcgt tatttgtggt tggccctctt gaagaaacca 180
tgatattgcg tggtatgagg tatcacccgg tagatatcga gaacacagtg acgagatgtc 240
atcgatcaat ctgtgaaagt gcggtcttca cgatgacaaa cctacttgtg gtagcagtgg 300
agcttgatgc agatgaacgc gaggcacttg acgtggttcc gctggtgacg acatccgtac 360
tgaatgaaca gcaacttgtc gtaggggtgg tggtagtggt tgaccctggt gtagtcccga 420

			173/234			
tcaattctcg	cggagagaaa	caacggatgc	atctgaggga	cgggttcctg	ggggaccagt	480
tggatcctat	ctacgtggcg	tataatatgt	agacacctca	ctgcttagtt	tcgtaattga	540
attgtgtcgt	agtttttta	aatgacaatt	aatagacaag	tttgaaattg	actgtagcgc	600
taggtttagg	tataaactag	cgtttggtaa	ggcaattatg	acaggaatta	ctgtcacgcg	660
tgacgcgaga	ccgtcacttc	acacgcaaac	ctgtggtcgc	С		701
<210> 182						
<211> 701						

<211> 701 <212> DNA

<213> Millepora sp. (Hydrozoan)

<400> 182 teegttateg etaaacagat gaeegettea aegttaagtt gaeaacagga ageaegaegg 60 agactgcagt cccgtacgcg cgaacgggat acctgggatt tatcaagaga acagatttca 120 cgcagacagg tggagcccgg catgacgcgt tatttgtggt tggccctctt gaaqaaacca 180 tgatattgcg tggtatgagg tatcacccgg tagatatcga gaacacagtg acgagatgtc 240 atcgatcaat ctgtgaaagt gcggtcttca cgatgacaaa cctacttgtg gtagcagtgg 300 agettgatge agatgaacge gaggeacttg aegtggttee getggtgaeg acateegtae 360 tgtatgaaca gcaacttgtc gtaggggtgg tggtagtggt tgaccctggt gtagtcccga 420 tcaattctcg cggagagaaa caacggatgc atctgaggga cgggttcctg ggggaccagt 480 tggatcctat ctacgtggcg tataatatgt agacacctca ctgcttagtt tcgtaattga 540 attgtgtcgt agtttttta aatgacaatt aatagacaag tttgaaattg actgtagcgc 600 taggtttagg tataaactag cgtttggtaa ggcaattatg acaggaatta ctgtcacgcg 660 tgacgcgaga ccgtcacttc acacgcaaac ctgtggtcgc c 701

<210> 183

<211> 701

<212> DNA

<213> Pavona decussaca

<220>

<221> CDS

#### WO 02/070703 PCT/GB02/00928 196/234

190/

<222> (1)..(699)

	gtt					atg Met							tga		cag Gln 15	48
-	_	_		-	_	cag Gln		-	_	-	-		-			96
						att Ile										144
						gcc Ala						tga		tgc Cys		192
gta Val	tga				cgg Arg	tag		tcg Ser	-		_	tga	_	gat Asp	-	240
						gtg Val							acc Thr		_	288
tgg Trp	tag	-		-	-	atg Met	_	_					_			336
	cgc Arg		_	_		ccg Pro		_	_		-		ttg Leu	_	tag	384
Gly ggg						acc Thr							att Ile 130			432
	_				_	atc Ile	_		-				ggg Gly 145		-	480
						cgt Arg									tag	528
	cgt Arg		tga		gtg Val	tcg Ser	tag	ttt Phe		taa	_		att Ile		-	576
caa Gln	-	tga	aat Asn	tga	ctg Leu	tag	cgc Arg 180	tag	gtt Val	tag	gta Val	taa		agc Ser		624
tgg Trp	taa				-	agg Arg			-		-	tga	_	gag Glu		672

197/234

gtc act tta cac gca aac ctg tgg tcg cc Val Thr Leu His Ala Asn Leu Trp Ser 200 205

701

<210> 184

<211> 13

<212> PRT

<213> Pavona decussaca

<400> 184

Ser Val Ile Ala Lys Gln Met Thr Ala Ser Thr Leu Ser 1  $\phantom{\bigg|}$  5  $\phantom{\bigg|}$  10

<210> 185

<211> 46

<212> PRT

<213> Pavona decussaca

<400> 185

Thr Trp Asp Leu Ser Arg Glu Gln Ile Ser Arg Arg Gln Met Glu Pro

Gly Met Thr Arg Tyr Leu Trp Leu Ala Leu Leu Lys Lys Pro 35 40 45

<210> 186

<211> 4

<212> PRT

<213> Pavona decussaca

<400> 186

Tyr Cys Val Val

<210> 187

<211> 4

<212> PRT

<213> Pavona decussaca

<400> 187

Gly Ile Thr Arg

<210> 188

<211> 5

<212> PRT

<213> Pavona decussaca

<400> 188

Ile Ser Arg Thr Gln

<210> 189

<211> 14

<212> PRT

<213> Pavona decussaca

<400> 189 ·

Arg Asp Val Ile Asp Gln Ser Val Lys Val Arg Ser Ser Arg

<210> 190

<211> 5

<212> PRT

<213> Pavona decussaca

<400> 190

Gln Thr Tyr Leu Trp 1 5

<210> 191

<211> 17 .

<212> PRT

<213> Pavona decussaca

<400> 191

Gln Trp Ser Leu Met Gln Met Asn Ala Arg His Leu Thr Trp Phe Arg 1 5 10 15

Trp

<210> 192

<211> 4

<212> PRT

<213> Pavona decussaca

<400> 192

Arg His Pro Tyr

<210> 193

<211> 6

<212> PRT

<213> Pavona decussaca

<400> 193

Met Asn Ser Asn Leu Ser 1 5

<210> 194

<211> 5

<212> PRT

<213> Pavona decussaca

<400> 194

Trp Leu Thr Leu Ala 1 5

<210> 195

<211> 13

<212> PRT

<213> Pavona decussaca

<400> 195

Ser Arg Ser Ile Leu Ala Giu Arg Asn Asn Gly Cys Ile 1 5 10

<210> 196

<211> 23

<212> PRT

<213> Pavona decussaca

<400> 196

Gly Thr Gly Ser Trp Gly Thr Ser Trp Ile Leu Ser Thr Trp Arg Ile 1 5 10 15

Ile Cys Arg His Leu Thr Ala 20

<210> 197

<211> 7

<212> PRT

<213> Pavona decussaca

<400> 197

<210> 198

<211> 4

<212> PRT

<213> Pavona decussaca

<400> 198

Thr Ser Val Trp

<210> 199

<211> 10

<212> PRT

<213> Pavona decussaca

<400> 199

Gly Asn Tyr Asp Arg Asn Tyr Cys His Ala 1  $\phantom{\bigg|}$  5  $\phantom{\bigg|}$  10

<210> 200

<211> 12

<212> PRT

<213> Pavona decussaca

<400> 200

Arg Glu Thr Val Thr Leu His Ala Asn Leu Trp Ser 1  $\phantom{\bigg|}5\phantom{\bigg|}$ 

<210> 201

<211> 231

<212> PRT

<213> coral

#### 202/234

<400> 201

Ser Val Ile Ala Lys Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr 1 5 10 15

Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys Pro 20 25 30

Tyr Glu Gly Glu Gln Thr Val Arg Leu Ala Val Thr Lys Gly Gly Pro 35 40 45

Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Cys Gln Tyr Gly Ser 50 55 60

Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys Gln 65 70 75 80

Ser Phe Pro Gly Arg Tyr Thr Trp Glu Arg Ile Met Asn Phe Glu Asp 85 90 95

Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys 100 105 110

Phe Ile Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly 115 120 125

Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu Arg 130 135 140

Leu Phe Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu 145 150 155 160

Lys Leu Glu Gly Gly His Tyr Leu Cys Glu Phe Lys Ser Thr Tyr 165 170 175

Lys Ala Arg Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Asp Arg 180 185 190

Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu Gln 195 200 205

Arg Glu Ile Ser Ile Ala Arg Lys Pro Leu Val Ala Cys Cys Phe Phe 210 215 220

Arg Val Lys Ser Arg His Lys 225 230

<210> 202

<211> 235

<212> PRT

<213> coral

<400> 202

#### 203/234

Ser Val Ile Ala Lys Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr 1 5 10 15

Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Leu Pro 20 25 30

Tyr Glu Gly Gln Thr Val Arg Leu Ala Val Thr Lys Gly Gly Pro 35 40 45

Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Cys Gln Tyr Gly Ser 50 55 60

Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys Gln 65 70. 75 80

Ser Phe Pro Gly Arg Tyr Thr Trp Glu Arg Ile Met Asn Phe Glu Asp 85 90 95

Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys 100 105 110

Phe Ile Tyr His Val Lys Arg Ser Gly Leu Asn Phe Pro Pro Asn Gly 115 120 125

Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu Arg 130 135 140

Leu Phe Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu 145 150 155 160

Lys Leu Glu Gly Gly His Tyr Leu Cys Glu Phe Lys Ser Thr Tyr 165 170 175

Lys Ala Lys Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Asp Arg 180 185 190

Lys Leu Asp Val Thr Asn His Asn Leu Asp Tyr Thr Ser Val Glu Gln 195 200 205

Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala Cys Arg Phe Phe 210 215 220

Arg Val Lys Ser Arg His Lys Tyr Ala Val Ala 225 230 235

<210> 203

<211> 49

<212> DNA

<213> oligonucleotide

<400> 203

tgagagaact agtctcgagc tctagaacaa gcttttttt ttttttt

49

V	0 02/0/0/03	204/234	PC 1/GB02/009
<211>	41		
<212>	DNA	•	
<213>	oligonucleotide		
<400>		gttatogota aacagatgac c	41
<210>	205		
<211>			
<212>			
<213>	oligonucleotide		
<400> gggtta		cgaccacagg tttgcgtg	38
.07.0:	006		
<210>			
<211> <212>			
	oligonucleotide		
72137	Oligonacieoliae		
<400>	206		
	aagt gccacctg		18
<210>	207		
<211>	19		
<212>	DNA		
<213>	oligonucleotide		
<400>	207		. 10
gttctg	aggt cattactgg		19
<210>	208		
<211>	20		

PCT/GB02/00928

WO 02/070703

<212> DNA

WO 02/070703		PCT/GB02/00928
	205/234	

<213> oligonucleotide

<400> 208 tcagggtact	tggtgaatgg					20
<210> 209						
<211> 669						
<212> DNA						
<213> Acro	opora sp					
<400> 209 ggatccgtta	tcgctaaaca	gatgacctac	: aaggtttata	tgtcaggcac	ggtcaatgga	60
cactactttg	aggtcgaagg	cgatggaaaa	ggaaagcctt	acgagggga	gcagacggta	120
aagctcactg	tcaccaaggg	tggacctctg	ccatttgctt	gggatatttt	atcaccacag	180
tcacagtacg	gaagcatacc	attcaccaag	taccctgaag	acatcccgga	ctatgtaaag	240
cagtcattcc	cggagggata	tacatgggag	aggatcatga	actttgaaga	tggtgcagtg	300
tgtactgtca	gcaatgactc	cagcatccaa	ggcaactgtt	tcatctacca	tgtcaagttc	360
tctggtttga	actttcctcc	caatggacct	gttatgcaga	agaagacaca	gggctgggaa	420
cccaacactg	agcgtctctt	tgcacgagat	ggaatgctga	taggaaacaa	ctttatggct	480
ctgaagttag	aaggaggtgg	tcactatttg	tgtgaattca	aatctactta	caaggcaaag	540
aagcctgtga	agatgccagg	gtatcactat	gttgaccgca	aactggatgt	aaccaatcac	600
aacaaggatt	acacttccgt	tgagcagtgt	gaaatttcca	ttgcacgcaa	acctgtggtc	660
gccctgcag						669
<210> 210						
<211> 222						
<212> PRT						
<213> Acro	pora sp					
<400> 210						
Gly Ser Val	. Ile Ala Ly 5	ys Gln Met	Thr Tyr Lys 10	Val Tyr Met	Ser Gly 15	
Thr Val Asr	Gly His Ty 20	yr Phe Glu	Val Glu Gly 25	Asp Gly Lys	Gly Lys	

#### 206/234

				•											
Pro	Tyr	Glu 35	Gly	Glu	Gln	Thr	Val 40	Lys	Leu	Thr	Val	Thr 45	Lys	Gly	Gly
Pro	Leu 50	Pro	Phe	Ala	Trp	Asp 55	Ile	Leu	Ser	Pro	Gln 60	Ser	Gln	Tyr	Gly
Ser 65	Ile	Pro	Phe	Thr	Lys 70	Tyr	Pro	Glu	Asp	Ile 75	Pro	Asp	Tyr	Val	Lys 80
Gln	Ser	Phe	Pro	Glu 85	Gly	Tyr	Thr	Trp	Glu 90	Arg	Ile	Met	Asn	Phe 95	Glu
Asp	Gly	Ala	Val 100	Суѕ	Thr	Val	Ser	Asn 105	Asp	Ser	Ser	Ile	Gln 110	Gly	Asn
Cys	Phe	Ile 115	Tyr	His	Val	Lys	Phe 120	Ser	Gly	Leu	Asn	Phe 125	Pro	Pro	Asn
Gly	Pro 130	Val	Met	Gln	Lys	Lys 135	Thr	Gln	Gly	Trp	Glu 140	Pro	Asn	Thr	Glu
Arg 145	Leu	Phe	Ala	Arg	Asp 150	Gly	Met	Leu	Ile	Gly 155	Asn	Asn	Phe	Met	Leu 160
Lys	Leu	Glu	Gly	Gly 165	Gly	His	Tyr	Leu	Cys 170	Glu	Phe	Lys	Ser	Thr 175	Tyr
Lys	Ala	Lys	Lys 180	Pro	Val	Lys	Met	Pro 185	Gly	Tyr	His	Tyr	Val 190	Asp	Arg
Lys	Leu	Asp 195	Val	Thr	Asn	His	Asn 200	Lys	Asp	Tyr	Thr	Ser 205	Val	Glu	Gln
Cys	Glu 210	Ile	Ser	Ile	Ala	Arg 215	Ъуs	Pro	Va1	Val	Ala 220	Leu	Gln		
<210	)> 2	211													
<21	L> 6	569													
<212	2> [	ANC							•						
<213	3> I	Disco	osoma	a sp											
<400		211	- دمدا	- 2220	בח מר	ata a	rctac		artt+	ata	tata	ragge	rac o	iat ca	ataa

cactactttg aggtcgaagg cgatggaaaa ggaaagcctt acgaggggga gcagacggta 120
aggctgactg tcaccaaggg cggacctctg ccatttgctt gggatatttt atcaccacag 180
tcacagtacg gaagcatacc attcaccaag taccctgaag acatccctga ctatgtaaag 240
cagtcattcc cggagggata tacatgggag aggatcatga actttgaaga tggtgcagtg 300
tgtactgtca gcaatgattc cagcatccaa ggcaactgtt tcatctacca tgtcaagttc 360
tctggtttga actttcctcc caatggacct gttatgcaga agaagacaca gggctgggaa 420

#### WO 02/070703 PCT/GB02/00928 207/234

<210> 212

<211> 222

<212> PRT

<213> Discosoma sp

<400> 212

Gly Ser Val Ile Ala Lys Gln Met Thr Tyr Lys Val Tyr Met Ser Gly 1 5 10 15

Thr Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys 20 25 30

Pro Tyr Glu Gly Glu Gln Thr Val Arg Leu Thr Val Thr Lys Gly Gly 35 40 45

Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Ser Gln Tyr Gly 50 55

Ser Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys 65 70 75 . 80

Gln Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ile Met Asn Phe Glu 85 90 95

Asp Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn 100 105 110

Cys Phe Ile Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn 115 120 125

Gly Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu 130 135 140

Arg Leu Leu Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Leu 145 150 155 160

Lys Leu Glu Gly Gly His Tyr Leu Cys Glu Phe Lys Ser Thr Tyr
165 170 175

Lys Ala Arg Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Asp Arg 180 185 190

Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu Gln 195 200 205

208/234

Arg Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala Leu Gln 210 <210> 213 <211> 669 DNA <212> <213> Sinularia sp <400> 213 qgatccgtta tcgctaaaca gatgacctac aaggtttata tgtcaggcac ggtcaatgga 60 cactactttg aggtcgaagg cgatggaaaa ggaaagcctt acgaggggga gcagacggta 120 aagctcactg tcaccaaggg tggacctctg ccatttgctt gggatatttt atcaccacag 180 tcacagtacg gaagcatacc attcaccaag taccctgaag acatcccgga ctatgtaaag 240 cagtcattcc cggaggggta tacatgggag aggatcatga actttgaaga tggtgcagtg 300 tgtactgtca gcaatgactc cagcatccaa ggcaactgtt tcatctacca tgtcaagttc 360 tetggtttga acttteette caatggaeet gttatgeaga agaagaeaca gggetgggaa 420 cccaacactg agcgtctctt tgcacgagat ggaatgctga taggaaacaa ctttatggct 480 ctgaagttag aaggaggtgg tcactatttg tgtgaattca aatctactta caaggcaaag 540 aagcctgtga agatgccagg gtatcactat gttgaccgca aactggatgt aaccaatcac 600 660 aacaaggatt acacttccgt tgagcagtgt gaaatttcca ttgcacgcaa acctttggtc 669 gccctgcag <210> 214 <211> 223 <212> PRT <213> Sinularia sp <400> 214 Gly Ser Val Ile Ala Lys Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys Pro Tyr Glu Gly Glu Gln Thr Val Lys Leu Thr Val Thr Lys Gly Gly

40

#### 209/234

									,					
Pro Leu 50	Pro	Phe	Ala	Trp	Asp 55	Ile	Leu	Ser	Pro	Gln 60	Ser	Gln	Tyr	Gly
Ser Ile 65	Pro	Phe	Thr	Lys 70	Tyr	Pro	Glu	Asp	Ile 75	Pro	Asp	Tyr	Val	Lys 80
Gln Ser	Phe	Pro	Glu 85	Gly	Tyr	Thr	Trp	Glu 90	Arg	Ile	Met	Asn	Phe 95	Glu
Asp Gly	Ala	Val 100	Суѕ	Thr	Val	Ser	Asn 105	Asp	Ser	Ser	Ile	Gln 110	Gly	Asn
Cys Phe	Ile 115	Tyr	His	Val	Lys	Phe 120	Ser	Gly	Leu	Asn	Phe 125	Pro	Ser	Asn
Gly Pro	Val	Met	Gln	Lys	Lys 135	Thr	Gln	Gly	Trp	Glu 140	Pro	Asn	Thr	Glu
Arg Leu 145	Phe	Ala	Arg	Asp 150	Gly	Met	Leu	Ile	Gly 155	Asn	Asn	Phe	Met	Ala 160
Leu Lys	Leu	Glu	Gly 165	Gly	Gly	His	Tyr	Leu 170	Cys	Glu	Phe	Lys	Ser 175	Thr
Tyr Lys	Ala	Lys 180	Lys	Pro	Val	Lys	Met 185	Pro	Gly	Tyr	His	Туг 190	Va1	Asp
Arg Lys	Leu 195	Asp	Val	Thr	Asn	His 200	Asn	Lys	Asp	Tyr	Thr 205	Ser	Val	Glu
Gln Cys 210	Glu	Ile	Ser	Ile	Ala 215	Arg	Lys	Pro	Leu	Val 220	Ala	Leu	Gln	
<210>	215													
<211>	669													
<212>	DNA													
<213>	Tuba	stre	a sp											

<400> 215 60 ggatccgtta tcgctaaaca gatgacctac aaggtttata tgtcaggcac ggtcaatgga cactactttg aggtcgaagg cgatggaaaa ggaaagcctt acgaggggga gcagacggta 120 180 aagctcactg tcaccaaggg tggacctctg ccatttgctt gggatatttt atcaccacag tcacagtacg gaagcatacc attcaccaag taccctgaag acatcccgga ctatgtaaag 240 cagtcattcc cggagggata tacatgggag aggatcatga actttgaaga tggtgcagtg 300 360 tgtactgtca gcaatgactc cagcatccaa ggcaactgtt tcatctacca tgtcaagttc tctggtttga actttcctcc caatggacct gttatgcaga agaagacaca gggctgggaa 420 cccaacactg agcgtctctt tgcacgagat ggaatgctga taggaaacaa ctttatggct 480 ctgaagttag aaggaggtgg tcactatttg tgtgaattca aatctactta caaggcaaag 540

600

660 669

210/234

aago	ctgt	iga a	gato	gccaç	ıg gt	atca	ctat	gtt:	gaco	gca	aact	ggat	gt a	aacca	atcac
aaca	agga	att a	cact	tccg	ıt to	gagca	gtgt	gaa	attt	cca	ttgo	gcgc	caa a	accto	jtggtc
gccc	etgea	ag													
<210	)> 2	216								•					
<211	.> 2	223													
<212	2> 1	PRT													
<213	3> :	<b>F</b> ubas	strea	a sp											
<400	)> 2	216													
Gly 1	Ser	Val	Ile	Ala 5	Lys	Gln	Met	Thr	Tyr 10	Lys	Val	Tyr	Met	Ser 15	Gly
Thr	Val	Asn	Gly 20	His	Tyr	Phe	Glu	Val 25	Glu	Gly	Asp	Gly	Lys 30	Gly	Lys
Pro	Tyr	Glu 35	Gly	Glu	Gln	Thr	Val 40	Lys	Leu	Thr	Val	Thr 45	Lys	Gly	Gly
Pro	Leu 50	Pro	Phe	Ala	Trp	Asp 55	Ile	Leu	Ser	Pro	Gln 60	Ser	Gln	Tyr	Gly
Ser 65	Ile	Pro	Phe	Thr	Lys 70	Tyr	Pro	Glu	Asp	Ile 75	Pro	Asp	Tyr	Val	Lys 80
Gln	Ser	Phe	Pro	Glu 85	Gly	Tyr	Thr	Trp	Glu 90	Arg	Ile	Met	Asn	Phe 95	Glu
Asp	Gly	Ala	Val 100	Cys	Thr	Val	Ser	Asn 105	Asp	Ser	Ser	Ile	Gln 110	Gly	Asn
Cys	Phe	Ile 115	Tyr	His	Val	Lys	Phe 120	Ser	Gly	Leu	Asn	Phe 125	Pro	Pro	Asn
Gly	Pro 130	Val	Met	Gln	Lys	Lys 135	Thr	Gln	Gly	Txp	Glu 140	Pro	Asn	Thr	Glu
Arg 145	Leu	Ьџе	Ala	Arg	Asp 150	Gly	Met	Leu	Ile	Gly 155	Asn	Asn	Phe	Met	Ala 160
Leu	Lys	Leu	Glu	Gly 165	Gly	Gly	His	Tyr	Leu 170	Cys	Glu	Phe	Lys	Ser 175	Thr
Tyr	Lys	Ala	Lys 180	Lys	Pro	Val	Lys	Met 185	Pro	Gly	Tyr	His	Tyr 190	Val	Asp
Arg	Lys	Leu 195	Asp	Val	Thr	Asn	His 200	Asn	Lys	Asp	Tyr	Thr 205	Ser	Val	Glu
Gln	Cys 210	Glu	I1e	Ser	Ile	Ala 215	Arg	Lys	Pro	Val	Val 220	Ala	Leu	Gln	

# WO 02/070703 PCT/GB02/00928 211/234

<210> 217 <211> 669 <212> DNA <213> Discosoma sp

<400> 217 ggatccgtta tcgctaaaca gatgacctac aaggtttata tgtcaggcac ggtcaatgga 60 cactactttg aggtcgaagg cgatggaaaa ggaaagcctt acgaggggga gcagacggta 120 aggctggctg tcaccaaggg cggacctctg ccatttgctt gggatatttt atcaccacag 180 tgtcagtacg gaagcatacc attcaccaag taccctgaag acatccctga ctatgtaaag 240 cagtcattcc cggagggatt tacatgggag aggatcatga actttgaaga tggtgcagtg 300 tgtcctgtca gcaatgattc cagcatccaa ggcaactgtt tcatctacca tgtcaagttc 360 tctggtttga actttcctcc caatggacct gttatgcaga agaagacaca gggctgggaa 420 ccccactctg agcgtctctt tgcacgagac ggaatgctga taggaaacac ctttatgqct 480 ctgaagttag aaggaggcgg tcactatttg tgtgaattca aaactactta caaggcaaag 540 aagcctgtga agatgccagg gtatcattat gttgaccgca aactggatgt aatcaatcac 600 aacaaggatt acactteegt tgageagtgt gaaattteea ttgeaegeaa acetgtggte 660 gccctgcag 669

<210> 218

<211> 223

<212> PRT

<213> Discosoma sp

<400> 218

Gly Ser Val Ile Ala Lys Gln Met Thr Tyr Lys Val Tyr Met Ser Gly 1 5 10 15

Thr Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys 20 25 30

Pro Tyr Glu Gly Glu Gln Thr Val Arg Leu Ala Val Thr Lys Gly Gly 35 40 45

Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Cys Gln Tyr Gly 50 55 60

#### 212/234

Ser 65	Ile	Pro	Phe	Thr	Lys 70	Tyr	Pro	Glu	Asp	Ile 75	Pro	Asp	Tyr	Val	Lys 80	
Gln	Ser	Phe	Pro	Glu 85	Gly	Phe	Thr	Trp	Glu 90	Arg	Ile	Met	Asn	Phe 95	Glu	
Asp	Gly	Ala	Val 100	Cys	Pro	Val	Ser	Asn 105	Asp	Ser	Ser	Ile	Gln 110	Gly	Asn	
Cys	Phe	Ile 115	Tyr	His	Val	Lys	Phe 120	Ser	Gly	Leu	Asn	Phe 125	Pro	Pro	Asn	
Gly	Pro 130	Val	Met	Gln	Lys	Lys 135	Thr	Gln	Gly	Trp	Glu 140	Pro	His	Ser	Glu	
Arg 145	Leu	Phe	Ala	Arg	Asp 150	Gly	Met	Leu	Ile	Gly 155	Asn	Thr	Phe	Met	Ala 160	
Leu	Lys	Leu	Glu	Gly 165	Gly	Gly	His	Tyr	Leu 170	Суѕ	Glu	Phe	Lys	Thr 175	Thr	
Tyr	Lys	Ala	Lys 180	Lys	Pro	Val	Lys	Met 185	Pro	Gly	Tyr	His	Tyr 190	Va1	Asp	
Arg	Lys	Leu 195	Asp	Val	Ile	Asn	His 200	Asn	Lys	Asp	Tyr	Thr 205	Ser	Val	Glu	
Gln	Cys 210	Glu	Ile	Ser	Ile	Ala 215	Arg	Lys	Pro	Val	Val 220	Ala	Leu	Gln		
<21	0>	219														
<21	1>	555														
<21	2>	DNA														
<21	3>	Sinu	lari	a sp												
<40	0>	219														
acg	gtaa	ggc	tggc	tgtc	ac c	aagg	gcgg	a cc	tctg	ccat	ttg	cttg	gga	tatt	ttatca	60
cca	cagt	gtc	agta	cgga	ag c	atac	catt	c ac	caag	tacc	ttg	aaga	cat	ccct	gactat	120
gta	aagc	agt	catt	cccg	ga g	ggat	ttac	a tg	ggag	agga	tca	tgaa	ctt	tgaa	gatggt	180
gca	gtgt	gta	ctgt	cagc	aa t	gatt	ccag	c at	ccaa	ggca	act	gttt	cat	ctac	catgtc	240
aag	ttct	ctg	gttt	gaac	tt t	cctc	ccaa	t gg	acct	gtta	tgc	agaa	gaa	gaca	cagggc	300

tgggaaccca acactgagcg tctctttgca cgagatggaa tgctgatagg aaacaacttt

atggctctaa agttagaggg aggtggtcac tatttgtgtg aattcaaatc tacttacaag gcaaagaagc ctgtgaagat gccagggtat cactatgttg accgcaaact ggatgtaacc

aatcacaaca aggattacac ttccgttgag cagtgtgaaa tttccattgc acgcaaacct

ttggtcgccc tgcag

360

420

480

540

555

213/234

<210> 220

<211> 223

<212> PRT

<213> Sinularia sp

<400> 220

Gly Ser Val Ile Ala Lys Gln Met Thr Tyr Lys Val Tyr Met Ser Gly
1 5 10 15

Thr Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys 20 25 30

Pro Tyr Glu Gly Glu Gln Thr Val Arg Leu Ala Val Thr Lys Gly Gly 35 40 45

Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Cys Gln Tyr Gly 50 55 60

Ser Ile Pro Phe Thr Lys Tyr Leu Glu Asp Ile Pro Asp Tyr Val Lys 65 70 75 80

Gln Ser Phe Pro Glu Gly Phe Thr Trp Glu Arg Ile Met Asn Phe Glu 85 90 95

Asp Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn 100 105 110

Cys Phe Ile Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn 115 120 125

Gly Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu 130 135 140

Arg Leu Phe Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala 145 150 155 160

Leu Lys Leu Glu Gly Gly Gly His Tyr Leu Cys Glu Phe Lys Ser Thr 165 170 175

Tyr Lys Ala Lys Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Asp 180 185 190

Arg Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu 195 200 205

Gln Cys Glu Ile Ser Ile Ala Arg Lys Pro Leu Val Ala Leu Gln 210 215 220

<210> 221

<211> 669

<212> DNA

WO 02/070703 PCT/GB02/00928 214/234

#### <213> Tubastrea sp

<400> 221	actaaaca	gatgacetae	aaggtttata	tatazaasa	aataaataa	60
ggatccgtta tc	gccaaaca	gatgatttat	aaggtttata	cyccayycac	ggicaaigga	60
cactactttg ag	gtcgaagg	cgatggaaaa	ggaaagcctt	acgagggga	gcagacggta	120
aggctggctg tc	accaaggg	cggacctctg	ccatttgctt	gggatatttt	atcaccacag	180
tgtcagtacg ga	agcatacc	attcaccaag	taccctgaag	acatccctga	ctatgtaaag	240
cggtcattcc cg	gagggatt	tacatgggag	aggatcatga	actttgaaga	tggtgcagtg	300
tgtactgtca gc	aatgattc	cagcatccaa	ggcaactgtt	tcatctacca	tgtcaagttc	360
tctggtttga ac	tttcctcc	caatggacct	gttatgcaga	agaagacaca	gggctgggaa	420
ccccactctg ag	cgtctctt	tgcacgagac	ggaatgctga	taggaaacaa	ctttatggct	480
ctgaagttag aa	ggaggcgg	tcactatttg	tgtgaattca	aaactactta	caaggcaaag	540
aagcctgtga ag	atgccagg	gtatcattat	gttgaccgca	aactggatgt	aatcaatcac	600
aacaaggatt ac	acttccgt	tgagcagtgt	gaaatttcca	ttgcacgcaa	acctgtggtc	660
gccctgcag						669

<210> 222

<211> 223

<212> PRT

<213> Tubastrea sp

<400> 222

Gly Ser Val Ile Ala Lys Gln Met Thr Tyr Lys Val Tyr Met Ser Gly 1 5 10 15

Thr Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys 20 25 30

Pro Tyr Glu Gly Glu Gln Thr Val Arg Leu Ala Val Thr Lys Gly Gly 35 40 45

Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Cys Gln Tyr Gly 50 55 60

Ser Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys 65 70 75 80

Arg Ser Phe Pro Glu Gly Phe Thr Trp Glu Arg Ile Met Asn Phe Glu 85 90 95

215/234 Asp Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn 105 Cys Phe Ile Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro His Ser Glu Arg Leu Phe Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala 155 Leu Lys Leu Glu Gly Gly Gly His Tyr Leu Cys Glu Phe Lys Thr Thr Tyr Lys Ala Lys Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Asp Arg Lys Leu Asp Val Ile Asn His Asn Lys Asp Tyr Thr Ser Val Glu 200 Gln Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala Leu Gln 210 215 220 <210> 223 <211> 46 <212> DNA <213> oligonucleotide <400> 223 cagggcgcgc caaggagata taacaatggc ttcctcagtt ctttcc 46 <210> 224 <211> 33 <212> DNA <213> oligonucleotide <400> 224 cactggatcc gcattgcact cttccgccgt tgc 33

<210> 225

<211> 45

<212> DNA

<213> oligonucleotide

216/234 <400> 225 gcatggcgcg ccaaggagat ataacaatga agactaatct ttttc 45

<210> 226 <211> 34 <212> DNA <213> oligonucleotide <400> 226 gcatggatcc gaattcggcc gaggataatg atag 34 <210> 227 <211> 45 <212> DNA <213> oligonucleotide <400> 227 gcatggcgcg ccaaggagat ataacaatga agactaatct ttttc 45 <210> 228 <211> 34 <212> DNA <213> oligonucleotide <400> 228 gcatggatcc gaattcggcc gaggataatg atag 34 <210> 229 <211> 46 <212> DNA <213> oligonucleotide

<400> 229

46

gatcttaatt aaagctcatc atgctgcagg gcgaccacag gtttgc

217/234

<210>	230	
<211>	42	
<212>	DNA	
<213>	oligonucleotide	
<400>	230	40
gcarct	gcag gtcgccacca gtaaaggaga agaacttttc ac	42
<210>	231	
<211>	39	
<212>	DNA	
<213>	oligonucleotide	
<400>	231 aatt aattatttgt atagttcatc catgccatg	39
009	and the control of th	0,5
<210>	232	
<211>	55	
<212>	DNA	
<213>	oligonucleotide	
<400> cagggc	232 gogo caaggagata taacaatggg atcogttato gotaaacaga tgaco	55
<210>	233	
<211>	45	
<212>	DNA	
<213>	oligonucleotide	
<400> ggctct	233 agaa aggagatata caatgtccgt tatcgctaaa cagat	45
<b>-010</b> :		
<210>	234	
<211>	45	

218/234

<212> DNA

<213> oligonucleotide

<400> 234
ggctctagaa aggagatata caatgtccgt tatcgctaaa cagat

45

<210> 235

<211> 50

<212> DNA

<213> oligonucleotide

<400> 235
ggcaagcttt cagtggtggt ggtggtggtg ggcgaccaca ggtttgcgtg 50

<210> 236

<211> 221

<212> PRT

<213> coral

<400> 236

Met Ser Val Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Ser Gly 1 5 10 15

Thr Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys 20 25 30

Pro Tyr Glu Gly Glu Gln Thr Val Lys Leu Thr Val Thr Lys Gly Gly 35 40 45

Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Cys Gln Tyr Gly 50 55 60

Ser Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys 65 70 75 80

Gln Ser Phe Pro Glu Gly Phe Thr Trp Glu Arg Ile Met Asn Phe Glu 85 90 95

Asp Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn 100 105 110

Cys Phe Thr Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn 115 120 125

#### 219/234

Gly Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Ser Ser Glu 130 135 140

His Leu Phe Ala Arg Gly Gly Met Leu Ile Gly Asn Asn His Met Ala 145 150 155 160

Leu Lys Leu Glu Gly Gly His Tyr Leu Cys Glu Phe Lys Thr Thr 165 170 175

Tyr Lys Ala Lys Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Asp 180 185 190

Arg Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu 195 200 205

Gln Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 210 215 220

<210> 237

<211> 221

<212> PRT

<213> coral

<400> 237

Met Ser Val Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Ser Gly
1 5 10 15

Thr Val Asn Gly His Tyr Phe Glu Val Gln Gly Asp Gly Lys Gly Lys 20 25 30

Pro Tyr Glu Glu Glu Gln Thr Val Lys Leu Thr Val Thr Lys Gly Gly 35  $\phantom{-}40\phantom{+}45\phantom{+}$ 

Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Cys Gln Tyr Gly 50 55 60

Ser Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys 70 75 80

Gln Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ile Met Asn Phe Glu 85 90 95

Asp Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn 100 105 110

Cys Phe Ile Tyr Asn Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn 115 120 125

Gly Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu 130 135 140

Arg Leu Phe Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala 145 150 155 160

#### 220/234

Leu Lys Leu Glu Gly Gly Gly His Tyr Leu Cys Glu Phe Lys Ser Thr 165 170 175

Tyr Lys Ala Lys Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Asp 180 185 190

Arg Lys Leu Asp Val Thr Asn His Asn Ile Asp Tyr Thr Ser Val Glu 195 200 205

Gln Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 210 215 220

<210> 238

<211> 226

<212> PRT

<213> coral

<400> 238

Ser Val Ile Ala Lys Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr 1 5 10 15

Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys Pro 20 25 30

Tyr Glu Gly Glu Gln Thr Val Arg Leu Thr Val Thr Lys Gly Gly Pro 35 40 45

Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Ser Gln Tyr Gly Ser 50 55 60

Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys Gln 65 70 75 80

Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ile Met Asn Phe Glu Asp 85 90 95

Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys 100 105 110

Phe Ile Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly 115 120 125

Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu Arg 130 135 140

Leu Phe Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu 145 150 155 160

Lys Leu Glu Gly Gly His Tyr Leu Cys Glu Phe Lys Ser Thr Tyr 165 170 175

Lys Ala Lys Lys Pro Val Arg Met Pro Gly Tyr His Tyr Val Asp Arg 180 · 185 190

#### 221/234

Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu Gln 195 200 205

Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala Trp Cys Phe Phe 210 215 220

Arg Val 225

<210> 239

<211> 220

<212> PRT

<213> coral

<400> 239

Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys Pro 20 25 30

Tyr Glu Gly Glu Gln Thr Val Arg Leu Ala Val Thr Lys Gly Gly Pro 35 40 45

Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Cys Gln Tyr Gly Ser 50 55 60

Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys Gln 65 70 75 80

Ser Phe Pro Gly Arg Tyr Thr Trp Glu Arg Ile Met Asn Phe Glu Asp 85 90 95

Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys 100 105 110

Phe Ile Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly 115 120 125

Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu Arg 130 135 140

Leu Phe Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu 145 150 155 160

Lys Leu Glu Gly Gly Gly His Tyr Leu Cys Glu Phe Lys Ser Thr Tyr 165 170 175

Lys Ala Arg Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Asp Arg 180 185 190

Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu Gln 195 200 205

222/234 Arg Glu Ile Ser Ile Ala Arg Lys Pro Leu Val Ala 215 <210> 240 <211> 230 <212> PRT <213> coral <400> 240 Met Ser Cys Ser Lys Asn Val Ile Lys Glu Phe Met Arg Phe Lys Val Arg Met Glu Gly Thr Val Asn Gly His Glu Phe Glu Ile Lys Gly Glu Gly Glu Gly Arg Pro Tyr Glu Gly His Cys Ser Val Lys Leu Met Val Thr Lys Gly Gly Pro Leu Pro Phe Ala Phe Asp Ile Leu Ser Pro Gln Phe Gln Tyr Gly Ser Lys Val Tyr Val Lys His Pro Ala Asp Ile Pro 105

Asp Tyr Lys Lys Leu Ser Phe Pro Glu Gly Phe Lys Trp Glu Arg Val

Met Asn Phe Glu Asp Gly Gly Val Val Thr Val Ser Gln Asp Ser Ser

Leu Lys Asp Gly Cys Phe Ile Tyr Glu Val Lys Phe Ile Gly Val Asn

Phe Pro Ser Asp Gly Pro Val Met Gln Arg Arg Thr Arg Gly Trp Glu

Ala Ser Ser Glu Arg Leu Tyr Pro Arg Asp Gly Val Leu Lys Gly Asp

Ile His Met Ala Leu Arg Leu Glu Gly Gly Gly His Tyr Leu Val Glu 170

Phe Lys Ser Ile Tyr Met Val Lys Lys Pro Ser Val Gln Leu Pro Gly

Tyr Tyr Tyr Val Asp Ser Lys Leu Asp Met Thr Ser His Asn Glu Asp

Tyr Thr Val Val Glu Gln Tyr Glu Lys Thr Gln Gly Arg His His Pro 220

Phe Ile Lys Pro Leu Gln 230

<210> 241

<211> 225

<212> PRT

<213> coral

<400> 241

Met Arg Ser Ser Lys Asn Val Ile Lys Glu Phe Met Arg Phe Lys Val 1 5 10 15

Arg Met Glu Gly Thr Val Asn Gly His Glu Phe Glu Ile Glu Gly Glu 20 25 30

Gly Glu Gly Arg Pro Tyr Glu Gly His Asn Thr Val Lys Leu Lys Val 35 40 45

Thr Lys Gly Gly Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln 50 55 60

Phe Gln Tyr Gly Asn Lys Val Tyr Val Lys His Pro Ala Asp Ile Pro 65 70 75 80

Asp Tyr Lys Lys Leu Ser Phe Pro Glu Gly Phe Lys Trp Glu Arg Trp 85 90 95

Met Asn Phe Glu Asp Gly Gly Val Val Thr Val Thr Gln Asp Ser Ser 100 105 110

Leu Gln Asp Gly Cys Phe Ile Tyr Lys Val Lys Phe Ile Gly Val Asn 115 120 125

Phe Pro Ser Asp Gly Pro Val Met Gln Lys Lys Thr Met Gly Trp Glu 130 135 140

Ala Ser Thr Lys Arg Leu Tyr Pro Arg Asp Gly Val Leu Lys Gly Glu 145 150 155 160

Ile His Lys Ala Leu Lys Leu Lys Asp Gly Gly His Tyr Leu Val Glu 165 170 175

Phe Lys Ser Ile Tyr Met Ala Lys Lys Pro Val Gln Leu Pro Gly Tyr 180 185 190

Tyr Tyr Val Asp Ser Lys Leu Asp Ile Thr Ser His Asn Glu Asp Tyr 195 200 205

Thr Ile Val Glu Gln Tyr Glu Arg Thr Glu Gly Arg His His Leu Phe 210 215 220

Leu

225

<210> 242

<211> 230

<212> PRT

<213> coral

<400> 242

Met Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu Val l $\phantom{a}$  5

Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly Glu 20 25 30

Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile Cys 35 40 45

Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr Phe 50 55 60

Ser Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys Arg 65 70 75 80

His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu Arg 85 90 95

Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu Val 100 105 110

Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly Ile 115 120 . 125

Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr Asn 130 135 140

Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn Gly 145 150 155 160

Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser Val 165 170 175

Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly Pro 180 185 190

Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu Ser 195 200 205

Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe Val 210 215 220

Thr Ala Ala Gly Ile Thr 225 230

<210> 243

<211> 818

<212> DNA

<213> Aequorea victoria

<400> 243						
ggatccaagg	agatataaca	atgaagacta	atctttttct	ctttctcatc	ttttcacttc	60
tcctatcatt	atcctcggcc	gaattcagta	aaggagaaga	acttttcact	ggagttgtcc	120
caattcttgt	tgaattagat	ggtgatgtta	atgggcacaa	attttctgtc	agtggagagg	180
gtgaaggtga	tgcaacatac	ggaaaactta	cccttaaatt	tatttgcact	actggaaaac	240
tacctgttcc	atggccaaca	cttgtcacta	ctttctctta	tggtgttcaa	tgcttttcaa	300
gatacccaga	tcatatgaag	cggcacgact	tcttcaagag	cgccatgcct	gagggatacg	360
tgcaggagag	gaccatcttc	ttcaaggacg	acgggaacta	caagacacgt	gctgaagtca	420
agtttgaggg	agacaccctc	gtcaacagga	tcgagcttaa	gggaatcgat	ttcaaggagg	480
acggaaacat	cctcggccac	aagttggaat	acaactacaa	ctcccacaac	gtatacatca	540
tggcagacaa	acaaaagaat	ggaatcaaag	ttaacttcaa	aattagacac	aacattgaag	600
atggaagcgt	tcaactagca	gaccattatc	aacaaaatac	tccaattggc	gatggccctg	660
tccttttacc	agacaaccat	tacctgtcca	cacaatctgc	cctttcgaaa	gatcccaacg	720
aaaagagaga	ccacatggtc	cttcttgagt	ttgtaacagc	tgctgggatt	acacatggca	780
tggatgaact	atacaaacat	gatgagcttt	aagagctc			818
<210> 244						
<211> 263						
<212> PRT						

<213> Aequorea victoria

<400> 244

Met Lys Thr Asn Leu Phe Leu Phe Leu Ile Phe Ser Leu Leu Ser

Leu Ser Ser Ala Glu Phe Ser Lys Gly Glu Glu Leu Phe Thr Gly Val

Val Pro Ile Leu Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe

Ser Val Ser Gly Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr 50 55 60

Leu Lys Phe Ile Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr

Leu Val Thr Thr Phe Ser Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro

Asp His Met Lys Arg His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly 100 105 110

Tyr Val Gln Glu Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys 115 120 125

Thr Arg Ala Glu Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile 130 135 140

Glu Leu Lys Gly Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His 145 150 155 160

Lys Leu Glu Tyr Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp 165 170 175

Lys Gln Lys Asn Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile 180 185 190

Glu Asp Gly Ser Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro 195 200 205

Ile Gly Asp Gly Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr 210 215 220

Gln Ser Ala Leu Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val 225 230 235. 240

Leu Leu Glu Phe Val Thr Ala Ala Gly Ile Thr His Gly Met Asp Glu 245 250 255

Leu Tyr Lys His Asp Glu Leu 260

<210> 245

<211> 235

<212> PRT

<213> Acropora aspera

<400> 245

Ser Val Ile Ala Lys Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr 1 5 10

Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Pro
20 25 30

Tyr Glu Gly Glu Gln Thr Val Arg Leu Ala Val Thr Lys Gly Gly Pro 35 40 45

Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Cys Gln Tyr Gly Ser 50 55 60

Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys Gln 65 70 75 80

#### 227/234

Ser Phe	Pro	Gly	Arg 85	Tyr	Thr	Trp	Glu	Arg 90	Ile	Met	Asn	Phe	Glu 95	Asp
Gly Ala	Val	Cys 100	Thr	Val	Ser	Asn	Asp 105	Ser	Ser	Ile	Gln	G1y 110	Asn	Суѕ

- Phe Ile Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly 115 120 125
- Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu Arg 130 135 140
- Leu Phe Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu 145 150 155 160
- Lys Leu Glu Gly Gly His Tyr Leu Cys Glu Phe Lys Ser Thr Tyr 165 170 175
- Lys Ala Lys Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Asp Arg 180 185 190
- Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu Gln . 195 200 205
- Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala Cys Arg Phe Phe 210 215 220
- Arg Val Lys Ser Arg His Lys Val Ala Val Ala 225 230 235
- <210> 246
- <211> 232
- <212> PRT
- <213> Acropora aspera

<400> 246

- Met Ala Ser Phe Leu Lys Lys Thr Met Pro Phe Lys Thr Thr Ile Glu 1 5 10 15
- Gly Thr Val Asn Gly His Tyr Phe Lys Cys Thr Gly Lys Gly Glu Gly 20 25 30
- Asn Pro Phe Glu Gly Thr Gln Glu Met Lys Ile Glu Val Ile Glu Gly 35 40
- Gly Pro Leu Pro Phe Ala Phe His Ile Leu Ser Thr Ser Cys Met Tyr 50 60
- Gly Ser Lys Thr Phe Ile Lys Tyr Val Ser Gly Ile Pro Asp Tyr Phe 65 70 75 80
- Lys Gln Ser Phe Pro Glu Gly Phe Thr Trp Glu Arg Thr Thr Tyr 85 90 95

#### 228/234

Glu Asp Gly Gly Phe Leu Thr Ala His Gln Asp Thr Ser Leu Asp Gly 105 Asp Cys Leu Val Tyr Lys Val Lys Ile Leu Gly Asn Asn Phe Pro Ala 120 Asp Gly Pro Val Met Gln Asn Lys Ala Gly Arg Trp Glu Pro Ala Thr 135 Glu Ile Val Tyr Glu Val Asp Gly Val Leu Arg Gly Gln Ser Leu Met Ala Leu Lys Cys Pro Gly Gly Arg His Leu Thr Cys His Leu His Thr Thr Tyr Arg Ser Lys Lys Pro Ala Ser Ala Leu Lys Met Pro Gly Phe His Phe Glu Asp His Arg Ile Glu Ile Met Glu Glu Val Glu Lys Gly 200 Lys Cys Tyr Lys Gln Tyr Glu Ala Ala Val Gly Arg Tyr Cys Asp Ala 220 Ala Pro Ser Lys Leu Gly His Asn 225 230 <210> 247 <211> 51 <212> DNA <213> oligonucleotide <400> 247 cgcgccaagg agatataaca atgagaggat cgcatcacca tcaccatcac g 51 <210> 248 <211> 51 <212> DNA <213> oligonucleotide <400> 248 51 gatccgtgat ggtgatggtg atgcgatcct ctcattgtta tatctccttg g <210> 249 <211> 47 <212> DNA

WO 02/070703 PCT/GB02/00928

<213> oligonucleotide

<400> ctgatta	249 aatt aaagctcatc atgtttgtat agttcatcca tgccatg	47
<210>	250	
<211>	34	
<212>	DNA	
<213>	oligonucleotide	
<400> gtgtgt;	250 actg tcagccagga ttccagcatc caag	34
<210>	251	
<211>	32	
<212>	DNA	
<213>	oligonucleotide	
<400>	251 gcaa tgatatcagc atccaaggca ac	32
ctgtca	geaa tyatateage ateeaayyea ac	32
<210>	252	
<211>	44	
<212>	DNA	
<213>	oligonucleotide	
<400> ggatcc	252 atcg ccaccatgtc taaaggtgaa gaattattca ctgg	44
<210>	253	
<211>	34	
<212>	DNA	

<213> oligonucleotide

<400> 253

	230/234	
cagctgt	ttat ttgtacaatt catccatacc atgg	34
<210>	254	
<211>	41	
<212>	DNA	
<213>	oligonucleotide	
<400> cgggat	254 ccat cgccaccatg aggtcttcca agaatgttat c	41
<210>	255 .	
<211>	31	
<212>	DNA	
<213>	oligonucleotide	
<400> gaggat	255 ccgc ggccgctaaa ggaacagatg g	31
<210>	256	
<211>	38	
<212>	DNA	
<213>	oligonucleotide	
<400> gaagat	256 ctaa aacaatgagt gtgatcgcta cacaaatg	38
<210>	257	
<211>	35	
<212>	DNA	
<213>	oligonucleotide	
<400> tatcaa	257 atcg ccggcgtcag gcgaccacag gtttg	35
<2105	258	

WO 02/070703

PCT/GB02/00928

	W O 02/070703	231/234	1 (1/(1/00/2/00/2)
<211>	30		
<212>	DNA		
<213>	oligonucleotide		
	258 tgtgt tgtgacgcaa ctgc	caactcc	30
<210>	259		
<211>	39		
<212>	DNA		
<213>	oligonucleotide		
	259 cageg gatecettea attt	agaaag caattgttc	39
	260		
<211>			
<212>	DNA		
<213>	oligonucleotide		
	· 260 atata ttacgcacca tatt	:c	25
<210>			
<211>			
<212>			
<213>	oligonucleotide		
<400>	261		
	tgacg acattggtag tc		22
<210>	262		
<211>			
<212>			

PCT/GB02/00928

WO 02/070703

232/234

<213> coral

<220>

<221> misc\_feature

<222> (15)..(15)

<223> x = any amino acid

<400> 262

Ser Pro Pro Asp Tyr Thr Leu Glu Phe Pro Lys Lys Xaa Val Ala 1 5 10 15

<210> 263

<211> 15

<212> PRT

<213> coral

<400> 263

Ser Pro Pro Asp Tyr Thr Leu Glu Arg Pro Lys Lys Gly Val Ala 1 5 10 15

<210> 264

<211> 24

<212> PRT

<213> coral

<400> 264

Asp Ser Ser Pro Glu Ser Tyr Leu Lys Asn Gly Ile Ala Glu Glu Met 1 5 10 15

Lys Thr Asp Val Met Glu Gly Ile 20

<210> 265

<211> 22

<212> PRT

233/234

<213> coral

<400> 265

Ser Tyr Leu Pro Asn Gly Ile Ala Glu Glu Met Lys Thr Asp Leu Met 1 5 10 15

Glu Gly Ile Val Asn Gly

<210> 266

<211> 22

<212> PRT

<213> coral

<400> 266

Ser Leu Tyr Gln Asn Gly Ile Ala Glu Glu Met Lys Thr Asp Leu Met 1 5 10  $\cdot$  15

20

20

Glu Gly Ile Val Asn Gly 20

<210> 267

<211> 20

<212> DNA

<213> oligonucleotide

<400> 267 atggaaggga tagtcgatgg

<210> 268

<211> 20

<212> DNA

<213> oligonucleotide

<400> 268 atggaaggga ttgtcgatgg

<210> 269

<pre>&lt;211&gt; 20 &lt;212&gt; DNA &lt;213&gt; oligonucleotide  &lt;400&gt; 269 atggaaggga tcgtcgatgg</pre>		W O 02/0/0/03	234/234	1 (1/ (3)/2/ (0)/2
<pre>&lt;213&gt; oligonucleotide  &lt;400&gt; 269 atggaaggga tcgtcgatgg</pre>	<211>	20		
<pre>&lt;400&gt; 269 atggaaggga tcgtcgatgg 20  &lt;210&gt; 270 &lt;211&gt; 19 &lt;212&gt; DNA &lt;213&gt; oligonucleotide  &lt;400&gt; 270 cctcgacaat cccttccat 19  &lt;210&gt; 271 &lt;211&gt; 19 &lt;212&gt; DNA &lt;213&gt; oligonucleotide</pre>	<212>	DNA ·		
atggaaggga tcgtcgatgg 20  <210> 270  <211> 19  <212> DNA  <213> oligonucleotide  <400> 270 cctcgacaat ccettccat 19  <210> 271  <211> 19  <212> DNA  <213> oligonucleotide	<213>	oligonucleotide		
<pre>&lt;211&gt; 19 &lt;212&gt; DNA &lt;213&gt; oligonucleotide  &lt;400&gt; 270 cctcgacaat cccttccat     19  &lt;210&gt; 271 &lt;211&gt; 19 &lt;212&gt; DNA &lt;213&gt; oligonucleotide  &lt;400&gt; 271</pre>				20
<pre>&lt;212&gt; DNA &lt;213&gt; oligonucleotide  &lt;400&gt; 270 cctcgacaat cccttccat</pre>	<210>	270		
<pre>&lt;213&gt; oligonucleotide  &lt;400&gt; 270     cctcgacaat cccttccat  19  &lt;210&gt; 271 &lt;211&gt; 19  &lt;212&gt; DNA &lt;213&gt; oligonucleotide  &lt;400&gt; 271</pre>	<211>	19		
<400> 270 cctcgacaat cccttccat  19  <210> 271  <211> 19  <212> DNA  <213> oligonucleotide  <400> 271	<212>	DNA		
<pre>cctcgacaat cccttccat 19  &lt;210&gt; 271 &lt;211&gt; 19  &lt;212&gt; DNA &lt;213&gt; oligonucleotide  &lt;400&gt; 271</pre>	<213>	oligonucleotide		
<211> 19 <212> DNA <213> oligonucleotide  <400> 271				19
<212> DNA <213> oligonucleotide  <400> 271	<210>	271		
<213> oligonucleotide  <400> 271	<211>	19		
<400> 271	<212>	DNA		
	<213>	oligonucleotide		
				19

WO 02/070703

PCT/GB02/00928

WO 02/070703	PCT/GB02/00928

135/234

agc ata co Ser Ile Pr 65	a ttc acc	aag tac Lys Tyr 70	cct gaa Pro Glu	gac Asp	atc Ile 75	cct Pro	gac Asp	tat Tyr	gta Val	aag Lys 80	240
cag tca tt Gln Ser Ph	c ccg gag e Pro Glu 85	gga ttt Gly Phe	aca tgg Thr Trp	gag Glu 90	agg Arg	atc Ile	atg Met	aac Asn	ttt Phe 95	gaa Glu	288
gat ggt gc Asp Gly Al	a gtg tgt a Val Cys 100	act gtc Thr Val	agc aat Ser Asn 105	gat Asp	tcc Ser	agc Ser	atc Ile	caa Gln 110	ggc Gly	aac Asn	336
tgt ttc ac Cys Phe Th	r Tyr His	gtc aag Val Lys	ttc tct Phe Ser 120	ggt Gly	ttg Leu	aac Asn	ttt Phe 125	cct Pro	ccc Pro	aat Asn	384
gga cct gt Gly Pro Va 130	g atg cag l Met Gln	aag aag Lys Lys 135	aca cag Thr Gln	ggc Gly	Trp	gaa Glu 140	ccc Pro	cac His	tct Ser	gag Glu	432
cgt ctc tt Arg Leu Ph 145	t gca cgg e Ala Arg	ggt gga Gly Gly 150	atg ctg Met Leu	ata Ile	gga Gly 155	aac Asn	aac Asn	ttt Phe	atg Met	gct Ala 160	480
ctg aag tt Leu Lys Le	a gaa gga u Glu Gly 165	Gly Gly	cac tat His Tyr	ttg Leu 170	tgt Cys	gaa Glu	ttc Phe	aaa Lys	act Thr 175	act Thr	528
tac aag go Tyr Lys Al	a aag aag a Lys Lys 180	cct gtg Pro Val	aag atg Lys Met 185	cca Pro	GJA aaa	tat Tyr	cat His	tat Tyr 190	gtt Val	gac Asp	576
cgc aaa ct Arg Lys Le 19	u Asp Val	acc aat Thr Asn	cac aac His Asn 200	aag Lys	gat Asp	tac Tyr	act Thr 205	tcc Ser	gtt Val	gag Glu	624
cag tgt ga Gln Cys Gl 210					Vaĺ '						663
<210> 129											
<211> 221											

<212> PRT

1.5

<213> Millepora sp.

<400> 129

Met Ser Val Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Ser Gly 1 5 10 15

Thr Val Asp Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys 20 25 30

#### 136/234

Pro Tyr Glu Gly Glu Gln Thr Val Lys Leu Thr Val Thr Lys Gly Gly 35 40 45

Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Cys Gln Tyr Gly 50 55 60

Ser Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys 65 70 75 80

Gln Ser Phe Pro Glu Gly Phe Thr Trp Glu Arg Ile Met Asn Phe Glu 85 90 95

Asp Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn 100 105 110

Cys Phe Thr Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn 115 120 125

Gly Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro His Ser Glu 130 135 140

Arg Leu Phe Ala Arg Gly Gly Met Leu Ile Gly Asn Asn Phe Met Ala 145 150 155 160

Leu Lys Leu Glu Gly Gly His Tyr Leu Cys Glu Phe Lys Thr Thr 165 170 175

Tyr Lys Ala Lys Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Asp 180 185 190

Arg Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu 195 200 205

Gln Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 210 215 220

<210> 130

<211> 663

<212> DNA

<213> Millepora sp.

## 137/234

<221> CDS

<222> (1)..(663)

Met	agt	130 gtg Val	atc Ile	gct Ala	aca Thr	caa Gln	atg Met	acc Thr	tac Tyr	aag Lys	gtt Val	tat Tyr	atg Met	tca Ser	ggc Gly	48
acg	gtc	gat	gga	5 cac	tac	ttt	gag	gtc	10 gaa	ggc	gat	gga	aaa	15 gga	aag	96
			20					25				ĞÎy	30		_	
Pro	Tyr	gag Glu 35	Gly	gag Glu	Gln	Thr	gta Val 40	aag Lys	Leu	Thr	gtc Val	acc Thr 45	aag Lys	ggc Gly	gga Gly	144
cct Pro	ctg Leu 50	cca Pro	ttt Phe	gct Ala	tgg Trp	gat Asp 55	att Ile	tta Leu	tca Ser	cca Pro	cag Gln 60	tgt Cys	cag Gln	tac Tyr	gga Gly	192
agc Ser 65	ata Ile	cca Pro	ttc Phe	acc Thr	aag Lys 70	tac Tyr	cct Pro	gaa Glu	gac Asp	atc Ile 75	cct Pro	gac 'Asp	tat Tyr	gta Val	aag Lys 80	240
cag Gln	tca Ser	ttc Phe	ccg Pro	gag Glu 85	gga Gly	ttt Phe	aca Thr	tgg Trp	gag Glu 90	agg Arg	atc Ile	atg Met	aac Asn	ttt Phe 95	gaa Glu	288
gat Asp	ggt Gly	gca Ala	gtg Val 100	tgt Cys	act Thr	gtc Val	agc Ser	aat Asn 105	ggt Gly	tcc Ser	agc Ser	atc Ile	caa Gln 110	ggc Gly	aac Asn	336
tgt Cys	ttc Phe	acc Thr 115	tac Tyr	cat His	gtc Val	aag Lys	ttc Phe 120	tct Ser	ggt Gly	ttg Leu	aac Asn	ttt Phe 125	cct Pro	ccc Pro	aat Asn	384
Gly	Pro 130	Val	Met	Gln	Lys	Lys 135	Thr	Gln	Gly	Trp	Glu 140	ccc Pro	His	Ser	Glu	432
cgt Arg 145	ctc Leu	ttt Phe	gca Ala	cgg Arg	ggt Gly 150	gga Gly	atg Met	ctg Leu	ata Ile	gga Gly 155	aac Asn	aac Asn	ttt Phe	atg Met	gct Ala 160	480
ctg Leu	aag Lys	tta Leu	gga Gly	gga Gly 165	Gly	Gly	His	Tyr	Leu	Cys	Glu	ttc Phe	Lys	act Thr 175	Thr	528
tac Tyr	agg Arg	gca Ala	aag Lys 180	aag Lys	cct Pro	gtg Val	aag Lys	atg Met 185	cca Pro	ggg Gly	tat Tyr	cat His	tat Tyr 190	gtt Val	gac Asp	576
cgc Arg	aaa Lys	ctg Leu 195	gat Asp	gta Val	acc Thr	aat Asn	cac His 200	aac Asn	aag Lys	gat Asp	tac Tyr	act Thr 205	tcc Ser	gtt Val	gag Glu	624

138/234

cag tgt gaa att tcc att gca cgc aaa cct gtg gtc gcc
Gln Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala
210

663

<210> 131

<211> 221

<212> PRT

<213> Millepora sp.

<400> 131

Met Ser Val Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Ser Gly
1 10 15

Thr Val Asp Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys 20 25 30

Pro Tyr Glu Gly Glu Gln Thr Val Lys Leu Thr Val Thr Lys Gly Gly 35 40 45

Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Cys Gln Tyr Gly 50 55 60

Ser Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys 65 70 75 80

Gln Ser Phe Pro Glu Gly Phe Thr Trp Glu Arg Ile Met Asn Phe Glu 85 90 95

Asp Gly Ala Val Cys Thr Val Ser Asn Gly Ser Ser Ile Gln Gly Asn 100 105 110

Cys Phe Thr Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn 115 120 125

Gly Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro His Ser Glu 130 140

Arg Leu Phe Ala Arg Gly Gly Met Leu Ile Gly Asn Asn Phe Met Ala 145 150 155 160

#### 139/234

Tyr Arg Ala Lys Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Asp 180 185 190

Arg Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu 195 200 205

Gln Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 210 215 220

<210> 132

<211> 660

<212> DNA

<213> Millepora sp.

<220>

<221> CDS

<400> 132

<222> (1)..(660)

100

agt Ser 1	gtg Val	atc Ile	gct Ala	aca Thr 5	caa Gln	atg Met	acc Thr	tac Tyr	aag Lys 10	gtt Val	tat Tyr	atg Met	tca Ser	ggc Gly 15	acg Thr	48	
gtc Val	gat Asp	gga Gly	cac His 20	tac Tyr	ttt Phe	gag Glu	gtc Val	gaa Glu 25	ggc Gly	gat Asp	gga G1y	aaa Lys	gga Gly 30	aag Lys	cct Pro	96	
tac Tyr	gag Glu	35 35	gag Glu	cag Gln	acg Thr	gta Val	aag Lys 40	ctc Leu	act Thr	gtc Val	acc Thr	aag Lys 45	ggc Gly	gga Gly	cct Pro	144	
ctg Leu	cca Pro 50	ttt Phe	gct Ala	tgg Trp	gat Asp	att Ile 55	tta Leu	tca Ser	cca Pro	cag Gln	tgt Cys 60	cag Gln	tac Tyr	gga Gly	agc Ser	192	
					tac Tyr 70											240	
					ttt											288	

90

336

Ser Phe Pro Glu Gly Phe Thr Trp Glu Arg Ile Met Asn Phe Glu Asp

ggt gca gtg tgt act gtc agc aat gat tcc agc atc caa ggc aac tgt

Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys

105

140/234											
ttc acc tac Phe Thr Tyr 115	His Val	aag ttc tct Lys Phe Ser 120	ggt ttg aac Gly Leu Asn	ttt cct ccc aat Phe Pro Pro Asn 125	gga 384 Gly						
cct gtg ato Pro Val Met 130	g cag aag : Gln Lys	aag aca cag Lys Thr Gln 135	ggc tgg gaa Gly Trp Glu	ccc cac tct gag Pro His Ser Glu 140	cgt 432 Arg						
ctc ttt gca Leu Phe Ala 145	cgg ggt Arg Gly	gga atg ctg Gly Met Leu 150	ata gga aac Ile Gly Asn 155	aac ttt atg gct Asn Phe Met Ala	ctg 480 Leu 160						
aag tta gaa Lys Leu Glu	gga ggc Gly Gly 165	ggt cac tat Gly His Tyr	ttg tgt gaa Leu Cys Glu 170	ttc aaa act act Phe Lys Thr Thr 175	tac 528 Tyr						
aag gca aag Lys Ala Lys	aag cct Lys Pro 180	gtg aag atg Val Lys Met	cca ggg tat Pro Gly Tyr 185	cat tat gtt gac His Tyr Val Asp 190	cgc 576 Arg						
	Val Thr			act tcc gtt gag Thr Ser Val Glu 205							
			cct gtg gtc Pro Val Val		660						
<210> 133											
<211> 220											
<212> PRT											
<213> Mill	epora sp.										
<400> 133											
Ser Val Ile 1	Ala Thr 5	Gln Met Thr	Tyr Lys Val	Tyr Met Ser Gly 15	Thr						
Val Asp Gly	His Tyr 20	Phe Glu Val	Glu Gly Asp 25	Gly Lys Gly Lys	Pro						
Tyr Glu Gly 35	Glu Gln	Thr Val Lys 40	Leu Thr Val	Thr Lys Gly Gly 45	Pro						

Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Cys Gln Tyr Gly Ser 50 60

Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys Gln 65 70 75 80

#### 141/234

Ser Phe Pro Glu Gly Phe Thr Trp Glu Arg Ile Met Asn Phe Glu Asp 85 90 95

Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys 100 105 110

Phe Thr Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly 115 120 125

Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro His Ser Glu Arg 130 135 140

Leu Phe Ala Arg Gly Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu 145 150 155

Lys Leu Glu Gly Gly His Tyr Leu Cys Glu Phe Lys Thr Thr Tyr 165 170 175

Lys Ala Lys Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Asp Arg 180 185 190

Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu Gln 195 200 205

Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 210 215 220

<210> 134

<211> 663

<212> DNA

<213> Porites murrayensis

<220>

<221> CDS

<222> (1)..(663)

<400> 134

atg agt gtg atc gct aca caa atg acc tac aag gtt tat atg cca ggc Met Ser Val Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Pro Gly 1 5 10 15

### 142/234

acg Thr	gtc Val	aat Asn	gga Gly 20	cac His	tac Tyr	ttt Phe	gag Glu	gtt Val 25	gaa Glu	ggc Gly	gat Asp	gga Gly	aaa Lys 30	gga Gly	aag Lys	96
cct Pro	tac Tyr	gag Glu 35	Gly	gag Glu	cag Gln	acg Thr	gta Val 40	aag Lys	ctc Leu	act Thr	gtc Val	acc Thr 45	aag Lys	ggc Gly	gga Gly	144
cct Pro	ctg Leu 50	cca Pro	ttt Phe	gct Ala	tgg Trp	gat Asp 55	att Ile	cta Leu	tca Ser	cca Pro	cag Gln 60	agt Ser	cag Gln	tac Tyr	gga Gly	192
agc Ser 65	ata Ile	cca Pro	ttc Phe	acc Thr	aag Lys 70	tac Tyr	cct Pro	gaa Glu	gac Asp	atc Ile 75	cct Pro	gac Asp	tat Tyr	gta Val	aag Lys 80	240
cag Gln	tca Ser	ttc Phe	cct Pro	gag Glu 85	gga Gly	tat Tyr	aca Thr	tgg Trp	gag Glu 90	agg Arg	atc Ile	atg Met	aac Asn	ttc Phe 95	gaa Glu	288
gat Asp	ggt Gly	gca Ala	gtg Val 100	tgt Cys	act Thr	gtc Val	agc Ser	aat Asn 105	gat Asp	tcc Ser	agc Ser	atc Ile	caa Gln 110	ggt Gly	aac Asn	336
tgt Cys	ttc Phe	atc Ile 115	tac Tyr	aat Asn	gtc Val	aag Lys	ttc Phe 120	tct Ser	ggt Gly	ttg Leu	aac Asn	ttt Phe 125	cct Pro	ccc Pro	aat Asn	384
gga Gly	cct Pro 130	gtt Val	atg Met	caa Gln	aag Lys	aag Lys 135	aca Thr	cag Gln	ggc Gly	tgg Trp	gaa Glu 140	ccc Pro	aac Asn	act Thr	gag Glu	432
cgt Arg 145	ctt Leu	tat Tyr	gca Ala	cga Arg	gat Asp 150	gga Gly	atg Met	ctg Leu	ata Ile	gga Gly 155	aac Asn	aac Asn	ttt Phe	atg Met	gct Ala 160	480
ctg Leu	aag Lys	ttg Leu	gaa G1u	gga Gly 165	ggt Gly	ggt Gly	cat His	tat Tyr	ttg Leu 170	tgt Cys	gaa Glu	ttc Phe	aaa Lys	tct Ser 175	act Thr	528
tac Tyr	aag Lys	gca Ala	aag Lys 180	aag Lys	cct Pro	gtg Val	atg Met	atg Met 185	cct Pro	gga Gly	tat Tyr	cac His	tat Tyr 190	gtt Val	gac Asp	576
cgc Arg	aaa Lys	ttg Leu 195	gat Asp	gta Val	acc Thr	aat Asn	cac His 200	aac Asn	aag Lys	gat Asp	tac Tyr	act Thr 205	tcc Ser	gtt Val	gag Glu	624
cag Gln	tgt Cys 210	gaa Glu	att Ile	tcc Ser	att Ile	gca Ala 215	cgc Arg	aaa Lys	cct Pro	gtg Val	gtc Val 220	gcc Ala				663

<210> 135

<211> 221

<212> PRT

<213> Porites murrayensis

<400> 135

Met Ser Val Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Pro Gly
1 5 10 15

Thr Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys 20 25 30

Pro Tyr Glu Gly Glu Gln Thr Val Lys Leu Thr Val Thr Lys Gly Gly 35 40 45

Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Ser Gln Tyr Gly 50 55 60

Ser Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys 65 70 75 80

Gln Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ile Met Asn Phe Glu 85 90 95

Asp Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn 100 105 110

Cys Phe Ile Tyr Asn Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn 115 120 125

Gly Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu 130 135 140

Arg Leu Tyr Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala 145 150 155 160

Leu Lys Leu Glu Gly Gly Gly His Tyr Leu Cys Glu Phe Lys Ser Thr 165 170 175

Tyr Lys Ala Lys Lys Pro Val Met Met Pro Gly Tyr His Tyr Val Asp 180 185 190

Arg Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu 195 200 205

Gln Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 210 215 220

· 144/234

144/104
<210> 136
<211> 663
<212> DNA
<213> Porites murrayensis
<220>
<221> CDS
<222> (1)(663)
<400> 136 atg agt gtg atc gct aca caa atg acc tac aag gtt tat atg tca ggc 48
Met Ser Val Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Ser Gly  1 5 10 15
acg gtc aat gga cac tac ttt gag gtt gaa ggc gat gga aaa gga aag 96
Thr Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys 20 25 30
cct tac gag ggg gag cag acg gta aag ctc act gtc acc aag ggc gga 144
Pro Tyr Glu Gly Glu Gln Thr Val Lys Leu Thr Val Thr Lys Gly Gly 35 40 45
cct ctg cca ttt gct tgg gat att cta tca cca cag agt cag tac gga 192
Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Ser Gln Tyr Gly 50 55 60
agc ata cca ttc acc aag tac cct gaa gac atc cct gac tat gta aag 240 Ser Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys
65 70 75 80
cag tca ttc cct gag gga tat aca tgg gag agg atc atg aac ttc gaa 288 Gln Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ile Met Asn Phe Glu
85 90 95
gat ggt gca gtg tgt act gtc agc aat gat tcc agc atc caa ggt aac Asp Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn
100 105 110
tgt ttc atc tac aat gtc aag ttc tct ggt ttg aac ttt cct ccc aat  Cys Phe Ile Tyr Asn Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn
115 120 125
gga cct gtt atg caa aag aag aca cag ggt tgg gaa ccc aac act gag Gly Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu
130 135 140
cgt ctc ttt gca cga gat gga atg ctg ata gga aac aac ttt atg gct 480 Arg Leu Phe Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala 145 150 155
145 150 155 160

### 145/234

									145	1234						
ctg Leu	aag Lys	ttg Leu	gaa Glu	gga G1y 165	ggt Gly	ggt Gly	cat His	tat Tyr	ttg Leu 170	tgt Cys	gaa Glu	ttc Phe	aaa Lys	tct Ser 175	act Thr	528
tac Tyr	aag Lys	gca Ala	aag Lys 180	aag Lys	cct Pro	gtg Val	atg Met	atg Met 185	cca Pro	Gly ggg	tat Tyr	cac His	tat Tyr 190	gtt Val	gac Asp	576
cgc Arg	aaa Lys	ttg Leu 195	gat Asp	gta Val	acc Thr	aat Asn	cac His 200	aac Asn	aag Lys	gat Asp	tac Tyr	act Thr 205	tcc Ser	gtt Val	gag Glu	624
cag Gln	tgt Cys 210	gaa Glu	att Ile	tcc Ser	att Ile	gca Ala 215	cgc Arg	aaa Lys	cct Pro	gtg Val	gtc Val 220	gcc Ala				663
<21	0>	137														
<21	1> :	221					•									
<21	2> :	PRT														
<21	3> :	Porit	tes n	nurra	ayens	sis										
< 40	0>	137														
Met 1	Ser	Val	Ile	Ala 5	Thr	Gln	Met	Thr	Tyr 10	Lys	Val	Tyr	Met	Ser 15	Gly	
Thr	Val	Asn	Gly 20	His	Tyr	Phe	Glu	Val 25	Glu	Gly	Asp	Gly	Lys 30	G1y	Lys	
Pro	Tyr	Glu 35	Gly	Glu	G1n	Thr	Val 40	Lys	Leu	Thr	Val	Thr 45	Lys	Gly	Gly	
Pro	Leu 50	Pro	Phe	Ala	Trp	Asp 55	Ile	Leu	Ser	Pro	Gln 60	Ser	Gln	Tyr	Gly	
Ser 65	Ile	Pro	Phe	Thr	Lys 70	Tyr	Pro	Glu	Asp	Ile 75	Pro	Asp	Tyr	Val	Lys 80	
Gln		Phe	Pro	Glu 85	Gly	Tyr	Thr	Trp ·	Glu 90	Arg	Ile	Met	Asn	Phe 95	Glu	
Asp			Val 100	Cys	Thr	Val	Ser	Asn 105	Asp	Ser	Ser	Ile	Gln 110	Gly	Asn	

Cys Phe Ile Tyr Asn Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn 115  $\phantom{\bigg|}$  120  $\phantom{\bigg|}$  125

### 146/234

130 135 140	
Arg Leu Phe Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala 145 150 155 160	
Leu Lys Leu Glu Gly Gly His Tyr Leu Cys Glu Phe Lys Ser Thr 165 170 175	
Tyr Lys Ala Lys Lys Pro Val Met Met Pro Gly Tyr His Tyr Val Asp 180 185 190	
Arg Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu 195 200 205	
Gln Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 210 215 220	
<210> 138	
<211> 660	
<212> DNA	
<213> Porites murrayensis	
•	
<220>	
<220>	
<220> <221> CDS <222> (1)(660)	
<220> <221> CDS	
<pre>&lt;220&gt; &lt;221&gt; CDS &lt;222&gt; (1)(660)  &lt;400&gt; 138 agt gtg atc gct aca caa atg acc tac aag gtt tat atg tca ggc acg Ser Val Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr</pre>	
<pre>&lt;220&gt; &lt;221&gt; CDS &lt;222&gt; (1)(660)  &lt;400&gt; 138 agt gtg atc gct aca caa atg acc tac aag gtt tat atg tca ggc acg Ser Val Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr 1</pre>	

									147	/234						10170	<b>D</b> 02,000
ata Ile 65	cca Pro	ttc Phe	acc Thr	aag Lys	tac Tyr 70	cct Pro	gac Asp	gac Asp	atc Ile	cct Pro 75	gac Asp	tat Tyr	gta Val	aaa Lys	cag Gln 80		240
tca Ser	ttc Phe	cct Pro	gag Glu	gga Gly 85	tat Tyr	aca Thr	tgg Trp	gag Glu	agg Arg 90	atc Ile	atg Met	aag Lys	ttt Phe	gaa Glu 95	gat Asp		288
ggt Gly	gca Ala	gtg Val	tgt Cys 100	act Thr	gtc Val	acc Thr	aat Asn	gac Asp 105	tcc Ser	agc Ser	atg Met	caa Gln	ggc Gly 110	aac Asn	tgt Cys		336
ttc Phe	atc Ile	tac Tyr 115	aat Asn	gtc Val	aag Lys	ttc Phe	tct Ser 120	ggt Gly	ttg Leu	aac Asn	ttt Phe	cct Pro 125	ccc Pro	aat Asn	gga Gly		384
cct Pro	gtt Val 130	atg Met	cag Gln	aag Lys	aag Lys	aca Thr 135	cag Gln	ggc Gly	tgg Trp	gaa Glu	ccc Pro 140	aac Asn	act Thr	GJ A aaa	cgt Arg		432
ctt Leu 145	tat Tyr	gca Ala	cga Arg	gat Asp	gga Gly 150	atg Met	ctg Leu	ata Ile	gga Gly	aac Asn 155	aac Asn	ttt Phe	atg Met	gct Ala	ctg Leu 160		480
aag Lys	ttg Leu	gaa Glu	gga Gly	ggt Gly 165	ggt Gly	cat His	tat Tyr	acc Thr	tgt Cys 170	gaa Glu	ttc Phe	aaa Lys	tct Ser	act Thr 175	tac Tyr		528
Lys	Ala	Lys	Lys 180	Pro	Val	atg Met	Met	Pro 185	Gly	Tyr	His	Tyr	Val 190	Asp	Arg		576
Lys	Leu	Asp 195	Val	Thr	Asn	cac His	Asn 200	Lys	Asp	Tyr	Thr	tcc Ser 205	gtt Val	gag Glu	cag Gln		624
tgt Cys	gaa Glu 210	att Ile	tcc Ser	att Ile	gca Ala	cgc Arg 215	aaa Lys	cct Pro	gtg Val	gtc Val	gcc Ala 220						660
<210	)> 1	.39															
<211	.> 2	20															
<212	:> P	RT															

<213> Porites murrayensis

<400> 139

Val Asn Gly His Tyr Phe Glu Val Gln Gly Asp Gly Lys Gly Lys Pro 20 25 30

148/234

Tyr Glu Gly Glu Gln Thr Val Lys Leu Thr Val Thr Lys Gly Gly Pro 35 40 45

Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Thr Gln Tyr Gly Ser 50 55 60

Ile Pro Phe Thr Lys Tyr Pro Asp Asp Ile Pro Asp Tyr Val Lys Gln 65 70 75 80

Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ile Met Lys Phe Glu Asp 85 90 95

Gly Ala Val Cys Thr Val Thr Asn Asp Ser Ser Met Gln Gly Asn Cys 100 105 110

Phe Ile Tyr Asn Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly 115 120 125

Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Gly Arg 130 135 140

Leu Tyr Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu 145 150 155 160

Lys Leu Glu Gly Gly His Tyr Thr Cys Glu Phe Lys Ser Thr Tyr 165 170 175

Lys Ala Lys Lys Pro Val Met Met Pro Gly Tyr His Tyr Val Asp Arg 180 185 190

Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu Gln 195 200 205

Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 210 215 220

<210> 140

<211> 660

<212> DNA

<213> Porites murrayensis

<220>

## 149/234

<221> CDS

<222> (1)..(660)

<400	)> ]	L40		•											
			gct Ala											-	48
_			cac His 20			_	_		_			_			96
			gag Glu											1	.44
			gct Ala											1	.92
			acc Thr											2	240
			gag Glu							_		_	-	. 2	88
	-		tgt Cys 100	-	-		-		-				_	3	36
			aat Asn											3	884
			caa Gln											4	32
			cga Arg											4	80
_	_	_	gga Gly	 			_	_	_					5	28
_	-	_	aag Lys 180		_	-					-	_	-	5	76
	_	_	gta Val				-	_			-		_	6	24

#### 150/234

tgt gaa att tcc att gca cgc aaa cct gtg gtc gcc
Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala
210 215 220

<210> 141

<211> 220

<212> PRT

<213> Porites murrayensis

<400> 141

Ser Val Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr 1 5 10

Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Gln Gly Lys Pro 20 25 30

Tyr Glu Gly Glu Gln Thr Val Lys Leu Thr Val Thr Lys Gly Gly Pro 35 40 45

Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Ser Gln Tyr Gly Ser 50 55 60

Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys Gln 65 70 75 80

Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ile Met Asn Phe Glu Asp  $85 \hspace{1.5cm} 90 \hspace{1.5cm} 95$ 

Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys 100 105 110

Phe Ile Tyr Asn Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly 115 120 125

Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu Arg 130 135

Leu Phe Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu 145 150 155 160

Lys Ser Glu Gly Gly His Tyr Leu Cys Glu Phe Lys Ser Thr Tyr 165 170 175

#### 151/234

Lys Ala Lys Lys Pro Val Met Met Pro Gly Tyr His Tyr Val Asp Arg 180 185 190

Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu Gln 195 200 205

Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 210 215 220

<210> 142

<211> 660

<212> DNA

<213> Porites murrayensis

<220>

<221> CDS

<222> (1)..(660)

<400> 142				
			gtt tat atg tca Val Tyr Met Ser	
			gat gga aaa gga Asp Gly Lys Gly 30	
		· • ·	gtc acc aag ggc Val Thr Lys Gly 45	J -
	Ala Trp Asp ]		cag agt cag tac Gln Ser Gln Tyr 60	
	-		cct gac tat gta Pro Asp Tyr Val 75	
			atc atg aac ttc Ile Met Asn Phe	
			agc atc caa ggc Ser Ile Gln Gly 110	

152/234

									132	234						
ttc Phe	atc Ile	tac Tyr 115	aat Asn	gtc Val	aag Lys	ttc Phe	tct Ser 120	ggt Gly	ttg Leu	aac Asn	ttt Phe	cct Pro 125	ccc Pro	aat Asn	gga Gly	384
cct Pro	gtt Val 130	atg Met	caa Gln	aag Lys	aag Lys	aca Thr 135	cag Gln	ggc Gly	tgg Trp	gaa Glu	ccc Pro 140	aac Asn	aca Thr	gag Glu	cgt Arg	432
ctc Leu 145	ttt Phe	gca Ala	cga Arg	gat Asp	gga Gly 150	atg Met	ctg Leu	ata Ile	gga Gly	aac Asn 155	aac Asn	ttt Phe	atg Met	gct Ala	ctg Leu 160	480
aag Lys	ttg Leu	gaa Glu	gga Gly	ggt Gly 165	ggt Gly	cat His	tat Tyr	ttg Leu	tgt Cys 170	gaa Glu	ttc Phe	aaa Lys	tct Ser	act Thr 175	tac Tyr	528
aag Lys	gca Ala	aag Lys	aag Lys 180	cct Pro	gtg Val	atg Met	atg Met	cca Pro 185	Gly Ggg	tat Tyr	cac His	tat Tyr	gtt Val 190	gac Asp	cgc Arg	576
aaa Lys	ttg Leu	gat Asp 195	gta Val	acc Thr	aat Asn	cac His	aac Asn 200	aag Lys	gat Asp	tac Tyr	act Thr	tcc Ser 205	gtt Val	gag Glu	cag Gln	624
					gca Ala											660
<210	)> ]	L43														
<211	.> 2	220														
<212	!> I	PRT														
<213	S> E	Porit	es n	nurra	yens	sis										
<400	)> ]	43													•	
Ser 1	Val	Ile	Ala	Thr 5		Met								Gly 15	Thr	·
Val	Asn	Gly	His 20	Tyr	Phe	Glu	Val	Glu 25	Gly	Asp	Gly	Lys	Gly 30	Lys	Pro .	
Tyr -	Glu	Gly 35	Glu	Gln	Thr	Val	Lys 40	Leu	Thr	Val	Thr	Lys 45	Gly	Gly	Pro	
Leu	Pro 50	Phe	Ala	Trp	Asp	Ile 55	Leu	Ser	Pro	Gln	Ser 60	Gln	Tyr	Gly	Ser	
Ile 65	Pro	Phe	Thr	Lys	Tyr 70	Pro	Glu	Asp	Ile	Pro 75	Asp	Tyr	Val	Lys	Gln 80	

#### 153/234

Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ile Met Asn Phe Glu Asp 85 90 95

Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys 100 105 110

Phe Ile Tyr Asn Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly 115 120 125

Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu Arg 130 135 140

Lys Leu Glu Gly Gly Gly His Tyr Leu Cys Glu Phe Lys Ser Thr Tyr 165 170 175

Lys Ala Lys Lys Pro Val Met Met Pro Gly Tyr His Tyr Val Asp Arg 180 185 190

Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu Gln 195 200 205

Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 210 215 220

<210> 144

<211> 660

<212> DNA

<213> Pink Pocillopora

<220>

<221> CDS

<222> (1)..(660)

48

WO 02/070703 PCT/GB02/00928													
WO 02/070703		154/234	PCT/GB02/00928										
gtc aat gga cac ta Val Asn Gly His Ty 20	c ttt gag gtt gaa r Phe Glu Val Glu 25	ggc gat gga aaa gga aa Gly Asp Gly Lys Gly Ly 30	g cct 96 s Pro										
tac gag ggc gag ca Tyr Glu Gly Glu Gl 35	g act gta aag ctc n Thr Val Lys Leu 40	act gtc acc aag ggc gg Thr Val Thr Lys Gly Gl 45	a cct 144 y Pro										
ctg ccg ttt gct tg Leu Pro Phe Ala Tr 50	g gat att tta tca p Asp Ile Leu Ser 55	cca cag act cag tac gg Pro Gln Thr Gln Tyr Gl 60	a agc 192 y Ser										
ata cca ttc acc aa Ile Pro Phe Thr Ly 65	g tac cct gaa gac s Tyr Pro Glu Asp 70	att cct gac tat gta aa Ile Pro Asp Tyr Val Ly 75	a cag 240 s Gln 80										
tca ttc cct gag gg Ser Phe Pro Glu Gl 85	a tat aca tgg gag y Tyr Thr Trp Glu	agg atc atg aag ttt ga Arg Ile Met Lys Phe Gl 90 95	a gat 288 u Asp										
ggt gca gta tgt ac Gly Ala Val Cys Th 100	t gtc agc aat gat r Val Ser Asn Asp 105	tcc agc atg caa ggc aa Ser Ser Met Gln Gly As: 110	c tgt 336 n Cys										
ttc atc tac aat gt Phe Ile Tyr Asn Va 115	c aag ttc tct ggt 1 Lys Phe Ser Gly 120	ttg aac ttt cct ccc aa Leu Asn Phe Pro Pro As 125	t gga 384 n Gly										
cct gtt atg cag aad Pro Val Met Gln Ly: 130	g aag aca cag ggc s Lys Thr Gln Gly 135	tgg gaa ccc aac act ga Trp Glu Pro Asn Thr Gl 140	g cgt 432 u Arg										
ctt tat gca cga ga Leu Tyr Ala Arg Asp 145	t gga atg ctg ata p Gly Met Leu Ile 150	gga aac aac ttt atg gc Gly Asn Asn Phe Met Ala 155	t ctg 480 a Leu 160										
	y Gly His Tyr Thr	tgt gaa ttc aaa tct ac Cys Glu Phe Lys Ser Th 170	r Tyr										
aag gca aag aag co Lys Ala Lys Lys Pro 180	t gtg atg atg cct o Val Met Met Pro 185	gga tat cac tat gtt ga Gly Tyr His Tyr Val Asp 190	c cgc 576 o Arg										
aaa ttg gat gta acc Lys Leu Asp Val Thi 195	c aat cac aac aag r Asn His Asn Lys 200	gat tac act tcc gtt gag Asp Tyr Thr Ser Val Glo 205	g cag 624 u Gln										
tgt gaa att tcc att Cys Glu Ile Ser Ile 210			660										
<210> 145													
<211> 220													
<212> PRT													

<212> PRT

<213> Pink Pocillopora

<400> 145

Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys Pro 20 25 30

Tyr Glu Gly Glu Gln Thr Val Lys Leu Thr Val Thr Lys Gly Gly Pro
35 40 45

Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Thr Gln Tyr Gly Ser 50 55 60

Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys Gln 65 70 75 80

Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ile Met Lys Phe Glu Asp 85 90 95

Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Met Gln Gly Asn Cys 100 105 110

Phe Ile Tyr Asn Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly 115 120 125

Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu Arg 130 135 140

Leu Tyr Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu 145 150 155

Lys Leu Glu Gly Gly His Tyr Thr Cys Glu Phe Lys Ser Thr Tyr 165 170 175

Lys Ala Lys Lys Pro Val Met Met Pro Gly Tyr His Tyr Val Asp Arg 180 185 190

Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu Gln 195 200 205

Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 210 215 220

		0 0 2, 0	,,,,,,,						156	/234						1 0 17 0 10 0 27
<210	> 1	146														
<211	> 6	663														
<212	> [	ANC														
<213	> I	Pink	Poci	i110p	ora											
												*				
<220	>															
<221	> (	CDS						•								
<222	<b>'&gt;</b>	(1)	. (663	3)												
<400 atg		gtg	atc	gct	aca	caa	atg	acc	tac	aag	gtt	tat	atg	tca	ggc	48
Met 1	Ser	Val	Ile	Ala 5	Thr	Gln	Met	Thr	Tyr 10	Lys	Val	Tyr	Met	Ser 15	Gly	
		aat														96
Thr	Val	Asn	20 GI y	His	Tyr	Phe	GLu	Val 25	Glu	СТĀ	Asp	GIA	30 Lys	GTÀ	Lys	
		gag Glu														144
110	1 y L	35	OLY	GIU	OIN	1111	40	Arg	neu	нα	var	45	шуЗ	GLY	Gry	
		cca Pro														192
	50				_	55					60	-		-	_	
		cca Pro														240
65					70					75					80	
_		ttc Phe	_	Glu					Glu			_		Phe	_	288
				85					90					95		. 226
		gca Ala														336
tat	ttc	atc		cat	atc	aad	ttc		aat	tta	aac	+++		ccc	aat	384
-		Ile 115			-	_		Ser		_						301
gga	cct	gtt	atg	cag	aag	aag			ggc	tgg	gaa	ccc	cac	tct	gag	432
Ğĺy	Pro 130	Val	Met	Gln	Lys	Lys 135	Thr	Gln	ĞÎУ	Trp	Glu 140	Pro	His	Ser	Glu	
		ttt														480
Arg 145	Leu	Phe	Ala	Arg	Asp 150	Gly	Met	Leu	Ile	Gly 155	Asn	Asn	Phe	Met	Ala 160	

## PCT/GB02/00928

# WO 02/070703 157/234 ctg aag tta gaa gga ggc ggt cac tat ttg tgt gaa ttc aaa act act 528 Leu Lys Leu Glu Gly Gly Gly His Tyr Leu Cys Glu Phe Lys Thr Thr tac aag gca aag act gtg aag atg cca ggg tat cat tat gtt gac 576 Tyr Lys Ala Lys Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Asp 185 cgc aaa ctg gat gta acc aat cac aac aag gat tac act tcc gtt gag 624 Arg Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu 200 cag tgt gaa att tcc att gca cgc aaa cct gtg gtc gcc 663 Gln Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 210 215 220 <210> 147 <211> 221 <212> PRT <213> Pink Pocillopora <400> 147 Met Ser Val Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Ser Gly 10 Thr Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys 20 Pro Tyr Glu Gly Glu Gln Thr Val Arg Leu Ala Val Thr Lys Gly Gly 35 40 Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Cys Gln Tyr Gly 50 55 Ser Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys Gln Ser Phe Pro Glu Gly Phe Thr Trp Glu Arg Ile Met Asn Phe Glu

Asp Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn 105

Cys Phe Ile Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn

### 158/234

100/100	
Gly Pro Val Met Gln Lys Lys Thr Gln Gly Trp 130 135	Glu Pro His Ser Glu 140
Arg Leu Phe Ala Arg Asp Gly Met Leu Ile Gly 145 150 155	Asn Asn Phe Met Ala 160
Leu Lys Leu Glu Gly Gly Gly His Tyr Leu Cys 165 170	Glu Phe Lys Thr Thr 175
Tyr Lys Ala Lys Lys Pro Val Lys Met Pro Gly 180 185	Tyr His Tyr Val Asp 190
Arg Lys Leu Asp Val Thr Asn His Asn Lys Asp 195 200	Tyr Thr Ser Val Glu 205
Gln Cys Glu Ile Ser Ile Ala Arg Lys Pro Val 210 215	Val Ala 220
<210> 148	
<212> DNA	
<213> Pink Pocillopora	
·	
<220>	
<221> CDS	
<222> (1)(663)	
<400> 148	
atg agt gtg atc gct aca caa atg acc tac aag	gtt tat atg tca ggc 48
Met Ser Val Ile Ala Thr Gln Met Thr Tyr Lys 1 10	Val Tyr Met Ser Gly 15
acg gtc aat gga cac tac ttt gag gtc gaa ggc	gat gga aaa gga aag 96
Thr Val Asn Gly His Tyr Phe Glu Val Glu Gly 20 25	Asp Gly Lys Gly Lys
cct tac gag ggg gag cag acg gta agg ctg gct Pro Tyr Glu Gly Glu Gln Thr Val Arg Leu Ala	Val Thr Lys Gly Gly
35 40	
	45
cct ctg cca ttt gct tgg gat att tta tca cca Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro	cag tgt cag tac gga 192

#### 159/234

agc Ser 65	ata Ile	cca Pro	ttc Phe	acc Thr	aag Lys 70	tac Tyr	cct Pro	gaa Glu	gac Asp	atc Ile 75	cct Pro	gac Asp	tat Tyr	gta Val	aag Lys 80	240
cag Gln	tca Ser	ttc Phe	ccg Pro	gag Glu 85	gga Gly	ttt Phe	aca Thr	tgg Trp	gag Glu 90	agg Arg	atc Ile	atg Met	aac Asn	ttt Phe 95	gaa Glu	288
gat Asp	ggt Gly	gca Ala	gtg Val 100	tgt Cys	act Thr	gtc Val	agc Ser	aat Asn 105	gat Asp	tcc Ser	agc Ser	atc Ile	caa Gln 110	Gly ggc	aac Asn	336
tgt Cys	ttc Phe	atc Ile 115	tac Tyr	cat His	gtc Val	aag Lys	ttc Phe 120	tct Ser	ggt Gly	ttg Leu	aac Asn	ttt Phe 125	cct Pro	ccc Pro	aat Asn	384
gga Gly	cct Pro 130	gtt Val	atg Met	cag Gln	aag Lys	aag Lys 135	aca Thr	cag Gln	ggc Gly	tgg Trp	gaa Glu 140	ccc Pro	cac His	tct Ser	gag Glu	432
cgt Arg 145	ctc Leu	ttt Phe	gca Ala	cga Arg	gat Asp 150	gga Gly	atg Met	ctg Leu	ata Ile	gga Gly 155	aac Asn	aac Asn	ttt Phe	atg Met	gct Ala 160	480
ctg Leu	aag Lys	tta Leu	gaa Glu	gga Gly 165	ggc Gly	ggt Gly	cac His	tat Tyr	ttg Leu 170	tgt Cys	gaa Glu	ttc Phe	aaa Lys	act Thr 175	act Thr	528
Tyr	Lys	Ala	Lys 180	Lys	Pro	gtg Val	Lys	Met 185	Pro	Gly	Tyr	His	Tyr 190	Val	Asp	576
cgc Arg	aaa Lys	ctg Leu 195	gat Asp	gta Val	acc Thr	aat Asn	cac His 200	aac Asn	aag Lys	gat Asp	tac Tyr	act Thr 205	tcc Ser	gtt Val	gag Glu	624
cag Gln	tgt Cys 210	gaa Glu	att Ile	tcc Ser	att Ile	gca Ala 215	cgc Arg	aaa Lys	cct Pro	gtg Val	gtc Val 220	gcc Ala	•		·	663

<210> 149

<211> 221

<212> PRT

<213> Pink Pocillopora

<400> 149

Met Ser Val Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Ser Gly 1 5 10 15

Thr Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys 20 25 30

#### 160/234

Pro Tyr Glu Gly Glu Gln Thr Val Arg Leu Ala Val Thr Lys Gly Gly 35 40 45

Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Cys Gln Tyr Gly 50 60

Ser Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys 65 70 75 80

Gln Ser Phe Pro Glu Gly Phe Thr Trp Glu Arg Ile Met Asn Phe Glu 85 90 95

Asp Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn 100 105 110

Cys Phe Ile Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn 115 120 125

Gly Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro His Ser Glu 130 135 140

Arg Leu Phe Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala 145 150 155 160

Leu Lys Leu Glu Gly Gly Gly His Tyr Leu Cys Glu Phe Lys Thr Thr 165 170 175

Tyr Lys Ala Lys Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Asp 180 185 190

Arg Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu 195 200 205

Gln Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 210 215 220

<210> 150

<211> 660

<212> DNA

<213> Pink Pocillopora

<220>

## 161/234

<221> CDS

<222> (1)..(660)

	<400> 150 agt gtg atc gct aca caa atg acc tac aag gtt tat atg tca ggc acg 48																
agt Ser 1	gtg Val	atc Ile	gct Ala	aca Thr 5	caa Gln	atg Met	acc Thr	tac Tyr	aag Lys 10	gtt Val	tat Tyr	atg Met	tca Ser	ggc Gly 15	acg Thr		48
gtc Val	aat Asn	gga Gly	cac His 20	tac Tyr	ttt Phe	gag Glu	gtt Val	gaa Glu 25	ggc Gly	gat Asp	gga Gly	aaa Lys	gga Gly 30	aag Lys	cct Pro		96
tac Tyr	gag Glu	ggg Gly 35	gag Glu	cag Gln	acg Thr	gta Val	aag Lys 40	ctc Leu	act Thr	gtc Val	acc Thr	aag Lys 45	ggc Gly	gga Gly	cct Pro		144
ctg Leu	cca Pro 50	ttt Phe	gct Ala	tgg Trp	gat Asp	att Ile 55	cta Leu	tca Ser	cca Pro	caġ Gln	agt Ser 60	cag Gln	tac Tyr	gga Gly	agc Ser		192
ata Ile 65	cca Pro	ttc Phe	acc Thr	aag Lys	tac Tyr 70	cct Pro	gaa Glu	gac Asp	atc Ile	cct Pro 75	gac Asp	tat Tyr	gta Val	aag Lys	cag Gln 80		240
tca Ser	ttc Phe	cct Pro	gag Glu	gga Gly 85	tat Tyr	aca Thr	tgg Trp	gag Glu	agg Arg 90	atc Ile	atg Met	aac Asn	ttc Phe	gaa Glu 95	gat Asp		288
ggt Gly	gca Ala	gtg Val	tgt Cys 100	act Thr	gtc Val	agc Ser	aat Asn	gat Asp 105	tcc Ser	agc Ser	atc Ile	caa Gln	ggc Gly 110	aac Asn	tgt Cys		336
ttc Phe	atc Ile	tac Tyr 115	aat Asn	gtc Val	aag Lys	ttc Phe	tct Ser 120	ggt Gly	ttg Leu	aac Asn	ttt Phe	cct Pro 125	ccc Pro	aat Asn	gga Gly		384
cct Pro	gtt Val 130	atg Met	caa Gln	aag Lys	aag Lys	aca Thr 135	cag Gln	ggc Gly	tgg Trp	gaa Glu	ccc Pro 140	aac Asn	act Thr	gag Glu	cgt Arg	٠	432
ctc Leu 145	ttt Phe	gca Ala	cga Arg	gat Asp	gga Gly 150	atg Met	ctg Leu	ata Ile	gga Gly	aac Asn 155	aac Asn	ttt Phe	atg Met	gct Ala	ctg Leu 160		480
aag Lys	ttg Leu	gaa Glu	gga Gly	ggt Gly 165	ggt Gly	cat His	tat Tyr	ttg Leu	tgt Cys 170	gaa Glu	ttc Phe	aaa Lys	tct Ser	act Thr 175	tac Tyr		528
				cct Pro													576
aaa Lys	ttg Leu	gat Asp 195	gta Val	acc Thr	aat Asn	cac His	aac Asn 200	aag Lys	gat Asp	tac Tyr	act Thr	tcc Ser 205	gtt Val	gag Glu	cag Gln		624

660

162/234

tgt gag att tcc att gca cgc aaa cct gtg gtc gcc Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 210 215 220

<210> 151

<211> 220

<212> PRT

<213> Pink Pocillopora

<400> 151

Ser Val Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr 1 5 10 15

Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys Pro 20 25 30

Tyr Glu Gly Glu Gln Thr Val Lys Leu Thr Val Thr Lys Gly Gly Pro 35 40 45

Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Ser Gln Tyr Gly Ser 50 55 60

Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys Gln 65 70 75 80

Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ile Met Asn Phe Glu Asp 85 90 95

Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys 100 105 110

Phe Ile Tyr Asn Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly 115 120 125

Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu Arg 130 135 140

Leu Phe Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu 145 150 155 160

Lys Leu Glu Gly Gly His Tyr Leu Cys Glu Phe Lys Ser Thr Tyr 165 170 175

### 163/234

Lys	Ala	Lys	Lys 180	Pro	Val	Met	Met	Pro 185	Gly	Tyr	His	Tyr	Val 190	Asp	Arg
Lys	Leu	Asp 195	Val	Thr	Asn	His	Asn 200	Lys	Asp	Tyr	Thr	Ser 205	Val	Glu	Gln

Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 210 215 220

<210> 152

<211> 663

<212> DNA

<213> Platygyra sp.

<220>

<221> CDS

<222> (1)..(663)

<400> 152								
			tac aag gtt ta Tyr Lys Val Ty 10					
acg gtc aa Thr Val As	gga cac ta Gly His Ty: 20	ttt gag gtc Phe Glu Val 25	gaa ggc gat gg Glu Gly Asp Gl	a aaa gga y Lys Gly 30	aag 96 Lys			
			ctc act gtc ac Leu Thr Val Th	r Lys Gly				
			tca cca cag tg Ser Pro Gln Cy 60					
			gac gtc cct ga Asp Val Pro As 75					
			gag agg atc at Glu Arg Ile Me 90					
			gat tcc agc at Asp Ser Ser Il					

164/234

		•	164/234						
tgt ttc acc Cys Phe Thr 115	tac cat g Tyr His V	atc aag ttc Val Lys Phe 120	tct ggt ttg Ser Gly Leu	aac ttt cct Asn Phe Pro 125	ccc aat Pro Asn	384			
gga cct gtg Gly Pro Val 130	atg cag a Mẹt Gln L	aag aag aca Lys Lys Thr 135	cag ggc tgg Gln Gly Trp	gaa ccc cac Glu Pro His 140	tct gag Ser Glu	432			
cgt ctc ttt Arg Leu Phe 145	Ala Arg G	ggt gga atg Gly Gly Met .50	ctg ata gga Leu Ile Gly 155	aac aac ttt Asn Asn Phe	atg gct Met Ala 160	480			
ctg aag tta Leu Lys Leu	gaa gga g Glu Gly G 165	ggc ggt cac Gly Gly His	tat ttg tgt Tyr Leu Cys 170	gga ttc aaa Gly Phe Lys	act act Thr Thr 175	528			
tac aag gca Tyr Lys Ala	aag aag c Lys Lys P 180	cct gtg aag Pro Val Lys	atg cca ggg Met Pro Gly 185	tat cat tat Tyr His Tyr 190	gtt gac Val Asp	576			
cgc aaa ctg Arg Lys Leu 195	gat gta a Asp Val T	cc aat cac Thr Asn His 200	aac aag gat Asn Lys Asp	tac att tcc Tyr Ile Ser 205	gtt gag Val Glu	624			
cag tgt gaa Gln Cys Glu 210	att tcc a Ile Ser I	itt gca cgc ile Ala Arg 215	aaa cct gtg Lys Pro Val	gtc gcc Val Ala 220		663			
<210> 153									
<211> 221									
<212> PRT									
<213> Platygyra sp.									
<400> 153									
Met Ser Val	Ile Ala T 5	hr Gln Met	Thr Tyr Lys	Val Tyr Met	Ser Gly 15				
Thr Val Asn	Gly His T 20	yr Phe Glu	Val Glu Gly 25	Asp Gly Lys	Gly Lys				
Pro Tyr Glu 35	Gly Glu A	arg Thr Val 40	Lys Leu Thr	Val Thr Lys 45	Gly Gly				
Pro Leu Pro 50	Phe Ala T	rp Asp Ile 55	Leu Ser Pro	Gln Cys Gln 60	Tyr Gly				
Asn Ile Pro 65		ys Tyr Pro O	Glu Asp Val 75	Pro Asp Tyr	Val Lys 80				

#### 165/234

Gln Ser Phe Pro Glu Gly Phe Thr Trp Glu Arg Ile Met Asn Phe Glu 85 90 95

Asp Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn 100 105 110

Cys Phe Thr Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn 115 120 125

Gly Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro His Ser Glu 130 135 140

Arg Leu Phe Ala Arg Gly Gly Met Leu Ile Gly Asn Asn Phe Met Ala 145 150 155 160

Leu Lys Leu Glu Gly Gly Gly His Tyr Leu Cys Gly Phe Lys Thr Thr 165 170 175

Tyr Lys Ala Lys Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Asp 180 185 190

Arg Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Ile Ser Val Glu 195 200 205

Gln Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 210 215 220

<210> 154

<211> 663

<212> DNA

<213> Platygyra sp.

<220>

<221> CDS

<222> (1)..(663)

<400> 154

atg agt gtg atc gct aca caa atg acc tac aag gtt tat atg tca ggc Met Ser Val Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Ser Gly 1 5 10 15

166/234 acg gtc aat gga cac tac ttt gag gtc gaa ggc gat gga aaa gga aag 96 Thr Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys 20 cct tac gag ggg gag cag acg gta agg ctc act gtc acc aag ggc gga 144 Pro Tyr Glu Gly Glu Gln Thr Val Arg Leu Thr Val Thr Lys Gly Gly cct ctg cca ttt gct tgg gat att ttg tca cca cag tat cag tac gga 192 Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Tyr Gln Tyr Gly age ata cea tte ace aag tae eet gaa gae ate eet gae tat gta aag 240 Ser Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys tag tca ttc ccg gag gga ttt aca tgg gac agg atc atg aac ttt gaa 288 Ser Phe Pro Glu Gly Phe Thr Trp Asp Arg Ile Met Asn Phe Glu gat ggt gca gtg tgt acc gtc agc aat gat tcc agc atc caa ggc aac 336 Asp Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn 105 tgt ttc atc tac cat gtc aag ttc tct ggt ttg aac ttt cct ccc aat 384 Cys Phe Ile Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn 115 120 gga cet gtt atg cag aag aag aca cag ggc tgg gaa cec aac act gag 432 Gly Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu cgt ctc ctt gca cga gat gga atg ctg cta gga aac aac ttt atg gct 480 Arg Leu Leu Ala Arg Asp Gly Met Leu Leu Gly Asn Asn Phe Met Ala 150 ctg aag tta gaa gga ggt ggt cac tat ttg tgt gaa ttc aaa act act 528 Leu Lys Leu Glu Gly Gly Gly His Tyr Leu Cys Glu Phe Lys Thr Thr 165 170 tac aag gca aag aag cct gtg aag atg cca ggg tat cac tat gtt gac 576 Tyr Lys Ala Lys Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Asp 180 cgc aaa ctg gat gta acc aat cac aac aag gat tac act tee gtt gag 624 Arg Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu

663

<210> 155

<211> 80

<212> PRT

<213> Platygyra sp.

210

cgg tgt gaa att tcc att gca cgc aaa cct gtg gtc gcc

Arg Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala

215

<400> 155

Met Ser Val Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Ser Gly  $1 \hspace{1cm} 5 \hspace{1cm} 10 \hspace{1cm} 15$ 

Thr Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys 20 25 30

Pro Tyr Glu Gly Glu Gln Thr Val Arg Leu Thr Val Thr Lys Gly Gly 35 40 45

Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Tyr Gln Tyr Gly 50 55 60

Ser Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys 65 70 75 80

<210> 156

<211> 140

<212> PRT

<213> Platygyra sp.

<400> 156

Ser Phe Pro Glu Gly Phe Thr Trp Asp Arg Ile Met Asn Phe Glu Asp 1 5 10 15

Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys 20 25 30

Phe Ile Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly 35 40 45

Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu Arg 50 55 60

Leu Leu Ala Arg Asp Gly Met Leu Leu Gly Asn Asn Phe Met Ala Leu 65 70 75 80

Lys Leu Glu Gly Gly His Tyr Leu Cys Glu Phe Lys Thr Thr Tyr 85 90 95

#### 168/234

Lys	Ala	Lys	Lys	Pro	Val	Lys	Met	Pro	Gly	Tyr	His	Tyr	Val	Asp	Arg
			100					105					110		_

Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu Arg 115 120 125

Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 130 135 140

<210> 157

<211> 660

<212> DNA

<213> Platygyra sp.

<220>

<221> CDS

<222> (1)..(660)

<400	)> 1	L <b>5</b> 7														
								tac Tyr								48
								gaa Glu 25								96
								ctc Leu								144
_			-		-			tca Ser		_	-	_				192
				_			_	gac Asp	-		-		_	_	-	240
								gag Glu								288
	_		-		-	-		gat Asp 105		-					_	336

169/234

										103/	234						
						aag Lys											384
						aag Lys											432
						gga Gly 150											480
						ggt Gly											528
						gtg Val											576
						aat Asn											624
						gca Ala											660
	<210	)> :	158													•	
	<211	i> 2	220														
	<212	2> 1	PRT												-		
٠	<213	3> 1	Platy	ygyra	a sp	•											٠
	<400	)> :	158														
	Ser 1	Val	Ile	Ala	Thr 5	Gln	Met					_			Gly 15		
	Val	Asn	Gly	His 20	Tyr	Phe	Glu	Val	Glu 25	Gly	Asp	Gly	Lys	Gly 30	Lys	Pro	
	Tyr	Glu	Gly 35	Glu	Gln	Thr	Val	Lys 40	Leu	Thr	Val	Thr	Lys 45	Gly	Gly	Pro	
	Leu	Pro 50	Phe	Ala	Trp	Asp	Ile 55	Leu	Ser	Pro	Gln	Cys 60	Gln	Tyr	Gly	Asn	
	Ile 65	Pro	Phe	Thr	Lys	Tyr 70	Pro	Glu	Asp	Val	Pro 75	Asp	Tyr	Val	Lys	Gln 80	

#### 170/234 .

Ser Phe Pro Glu Gly Phe Thr Trp Glu Arg Ile Met Asn Phe Glu Asp

Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys 100 105 110

Phe Thr Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly 115 120 125

Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro His Ser Glu Arg 130 135 140

Leu Phe Ala Arg Gly Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu 145 150 155 160

Lys Leu Glu Gly Gly His Tyr Leu Cys Gly Phe Lys Thr Thr Tyr 165 170 175

Lys Ala Lys Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Asp Arg 180 185 190

Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Ile Ser Val Glu Gln 195 200 205

Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 210 215 220

<210> 159

<211> 660

<212> DNA

<213> Platygyra sp.

<220>

<221> CDS

<222> (1)..(660)

<400> 159

agt gtg atc gct aca caa atg acc tac aag gtt tat atg tca ggc acg Ser Val Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr 1 5 10 15

#### 171/234

		gga Gly														96	5
		ggg Gly 35		_	_	_	-			-		_				144	
		ttt Phe														192	?
		ttc Phe		-			_	-	_		_		_	_	_	240	)
		ccg Pro														288	3
		gtg Val														336	5
		tac Tyr 115														384	1
		atg Met	_	_	-		_			_					_	432	2
		gca Ala														480	)
		gaa Glu														528	3
_	_	aag Lys	_			_	_						-	-	-	576	
		gat Asp 195											Val			624	4
_	_	att Ile			-	-				_	-					660	0

<210> 160

<211> 220

<212> PRT

<213> Platygyra sp.

<400> 160

Ser Val Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr 1 5 10 15

Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys Pro

Tyr Glu Gly Glu Gln Thr Val Lys Leu Thr Val Thr Lys Gly Gly Pro
35 40 45

Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Cys Gln Tyr Gly Asn 50 55 60

Ile Pro Phe Thr Lys Tyr Pro Glu Asp Val Pro Asp Tyr Val Lys Gln 65 70 75 80

Ser Phe Pro Glu Gly Phe Thr Trp Glu Gly Ile Met Asn Phe Glu Asp 85 90 95

Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys 100 105 110

Phe Thr Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly 115 120 125

Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro His Ser Glu Arg 130 135 140

Leu Phe Ala Arg Gly Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu 145 150 155 160

Lys Leu Glu Gly Gly His Tyr Leu Cys Gly Phe Lys Thr Thr Tyr 165 170 175

Lys Ala Lys Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Asp Arg 180 185 190

Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Ile Ser Val Glu Gln 195 200 205

Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 210 215 220

# 173/234

<21	0>	161											•			
<21	1>	660														
<21	2>	DNA														
<21	3>	Pavo	na d	ecus	sata											
<22	0>															
<22	1> (	CDS														
<22	2>	(1).	. (66	0)												
		ė														
<400 agt		161 atc	act	aca	caa	atα	acc	tac	aad	att	tat	ato	tca	aac	aca	. 48
Ser 1	Val	Ile	Ala	Thr 5	Gln	Met	Thr	Tyr	Lys 10	Val	Tyr	Met	Ser	Gly 15	Thr	
gtc	aat	gga	cac	tac	ttt	gag	gtt	gaa	ggc	gat	gga	aaa	gga	gag	cct	96
val	ASII	Gly	20	ıyr	Pile	GIU	vaı	25	стХ	Asp	стА	ьys	30	GIu	Pro	
tac	gag	Gly ggg	gag	cag	acg	gta Vəl	agg	ctc	act	gtc	aca	aag	ggc	gga	cct	144
-,-		35	024	<b></b>			40	Deu		vai	1111	45	G-Y	GLY	110	
ctg Leu	cca Pro	ttt Phe	gct Ala	tgg Trp	gat Asp	att Ile	tta Leu	tca Ser	cca Pro	cag Gln	tat Tyr	cag Gln	tac Tyr	gga Gly	agc Ser	192
	50					55					60					
Ile		ttc Phe			Tyr											240
65					70					75					80	
Ser	ttc Phe	ccg Pro	gaa Glu	Gly	tat Tyr	aca Thr	tgg Trp	gag Glu	Arg	atc Ile	atg Met	aac Asn	ttt Phe	Glu	gat Asp	288
				85					90					95		
Gly	Ala	gtg Val	Cys 100	Thr	Val	agc Ser	Asn	gat Asp 105	Ser	agc Ser	Ile	caa Gln	Gly	aac Asn	tgt Cys	336
ttc	atc	tac		ata	nee	++-	tet		++~	220	+++	cat	110	+	gga	384
Phe	Ile	Tyr 115	His	Val	Lys	Phe	Ser 120	Gly	Leu	Asn	Phe	Pro	Pro	Asn	Gly	304
cct	ata		caq	aaσ	aaα	aca		aac	taa	σaa	ccc	_	act	nan	cgt	. 432
Pro	Val 130	Thr	Gln	Lys	Lys	Thr 135	Gln	ĞÎy	Trp	Ğlu	Pro 140	Asn	Thr	Glu	Arg	2010
ctc	ttt	gca	cga	gat	gga	atg	ctg	ata	gga	aac	aac	ttt	atg	gct	ctg	480
Leu 145	Phe	Ala	Arg	Asp	Gly 150	Met	Leu	Ile	Gly	Asn 155	Asn	Phe	Met	Ala	Leu 160	

528

# aag tta gaa gga ggc ggt cac tat ttg tgt gaa ttc aaa tcg act tac Lys Leu Glu Gly Gly His Tyr Leu Cys Glu Phe Lys Ser Thr Tyr 165 170 175

aag gca aag aag act gtg aag atg cca ggg tat cac tat gtt gac cgc 576 Lys Ala Lys Lys Thr Val Lys Met Pro Gly Tyr His Tyr Val Asp Arg 180 185 190

aaa ctg gat gta acc aat cac aac aag gat tac act tcc gtt gag cag
Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu Gln
195 200 205

tgt gaa att tcc att gca cgc aaa cct gtg gtc gcc
Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala
210 215 220

<210> 162

<211> 220

<212> PRT

<213> Pavona decussata

<400> 162

Ser Val Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr 1 5 10 15

Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Glu Pro 20 25 30

Tyr Glu Gly Glu Gln Thr Val Arg Leu Thr Val Thr Lys Gly Gly Pro 35 40 45

Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Tyr Gln Tyr Gly Ser 50 55 60

Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys Gln 65 70 75 80

Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ile Met Asn Phe Glu Asp 85 90 95

Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys 100 105 110

Phe Ile Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly 115 120 125

#### 175/234

130 135 140	
Leu Phe Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu 145 150 155 160	
Lys Leu Glu Gly Gly His Tyr Leu Cys Glu Phe Lys Ser Thr Tyr 165 170 175	
Lys Ala Lys Lys Thr Val Lys Met Pro Gly Tyr His Tyr Val Asp Arg 180 185 190	
Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu Gln 195 200 205	
Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 210 215 220	
<210> 163	
<211> 663	
<212> DNA	
<213> Pavona decussata	
<220>	
<220>	
<220> <221> CDS <222> (1)(663)  <400> 163	
<220> <221> CDS <222> (1)(663)	48
<220> <221> CDS <222> (1)(663)  <400> 163 atg agt gtg atc gct aca caa atg acc tac aag gtt tat atg tca ggc Met Ser Val Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Ser Gly	48 96
<pre>&lt;220&gt; &lt;221&gt; CDS &lt;222&gt; (1)(663)  &lt;400&gt; 163 atg agt gtg atc gct aca caa atg acc tac aag gtt tat atg tca ggc Met Ser Val Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Ser Gly 1</pre>	•

#### 176/234

agc Ser 65	ata Ile	cca Pro	ttc Phe	acc Thr	aag Lys 70	tac Tyr	cct Pro	gaa Glu	gac Asp	atc Ile 75	cct Pro	gac Asp	tat Tyr	gta Val	tag	240
cag Gln 80	tca Ser	ttc Phe	ccg Pro	gaa Glu	gga Gly 85	tat Tyr	aca Thr	tgg Trp	gag Glu	agg Arg 90	atc Ile	atg Met	aac Asn	ttt Phe	gaa Glu 95	288
gat Asp	ggt Gly	gct Ala	gtg Val	tgt Cys 100	act Thr	gtc Val	agc Ser	aat Asn	gat Asp 105	tcc Ser	agc Ser	atc Ile	caa Gln	ggc Gly 110	aac Asn	336
tgt Cys	ttc Phe	atc Ile	tac Tyr 115	cat His	gtc Val	aag Lys	ttt Phe	tct Ser 120	ggt Gly	ttg Leu	aac Asn	ttt Phe	cct Pro 125	ccc Pro	aat Asn	384
gga Gly	cct Pro	gtg Val 130	atg Met	cag Gln	aag Lys	aag Lys	aca Thr 135	cag Gln	ggc Gly	tgg Trp	gaa Glu	ccc Pro 140	aac Asn	act Thr	gag Glu	432
					gat Asp											480
					ggc Gly 165											528
					act Thr											576
					acc Thr											624
					att Ile											663

<210> 164

<211> 79

<212> PRT

<213> Pavona decussata

<400> 164

Met Ser Val Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Ser Gly  $1 \hspace{1cm} 5 \hspace{1cm} 10 \hspace{1cm} 15$ 

Thr Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys 20 25 30

#### 177/234

Pro Tyr Glu Glu Glu Gln Thr Val Arg Leu Thr Val Thr Lys Gly Gly 35 40

Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Tyr Gln Tyr Gly 50 55

Ser Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val 65 70 75

<210> 165

<211> 141

<212> PRT

<213> Pavona decussata

<400> 165

Gln Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ile Met Asn Phe Glu 1 5 10 15

Asp Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn 20 25 30

Cys Phe Ile Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn 35 40 45

Gly Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu 50 60

Arg Leu Phe Ala Arg Asp Gly Leu Leu Ile Gly Asn Asn Phe Met Ala 65 70 75 80

Leu Lys Leu Glu Glu Gly Gly His Tyr Leu Cys Glu Phe Lys Ser Thr 85 90 95

Tyr Lys Ala Lys Lys Thr Ala Lys Met Pro Gly Tyr His Tyr Val Asp 100 105 110

Arg Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu 115 120 125

Gln Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 130 135 140

#### 178/234

<21	.0>	166															
<21	1>	663															
<21	.2>	DNA															٠
<21	3>	Pavo	na d	lecus	sata								-				
<22	0>																
<22	1>	CDS															
<22	2>	(1).	. (66	3)													
<40 atg		166 ata	atc	act	aca	caa	ata	acc	tac	aan	att	tst	240	+	~~~		40
Met 1	Ser	Vaĺ	Ile	Äla 5	Thr	Gln	Val	Thr	Tyr 10	Lys	Val	Tyr	Met	Ser 15	Gly		48
acg Thr	gtc Val	aat Asn	gga Gly 20	cac	tac Tyr	ttt Phe	gag Glu	gtt Val 25	gaa Glu	ggc Gly	gat Asp	gga Gly	aaa Lys 30	gga Gly	aag Lys		96
cct Pro	tac Tyr	gag Glu 35	Gly ggg	gag Glu	caa Gln	acg Thr	gta Val 40	agg Arg	ctc Leu	act Thr	gtc Val	aca Thr 45	aag Lys	ggc Gly	gga Gly	1	44
Pro	ctg Leu 50	cca Pro	ttt Phe	gct Ala	tgg Trp	gat Asp 55	att Ile	tta Leu	tca Ser	cca Pro	cag Gln 60	tat Tyr	cag Gln	tac Tyr	gga Gly	1	92
agc Ser 65	ata Ile	cca Pro	ttc Phe	acc Thr	aag Lys 70	tac Tyr	cct Pro	gaa Glu	gac Asp	atc Ile 75	cct Pro	gac Asp	tat Tyr	gta Val	aag Lys 80	2	40
cag Gln	tca Ser	ttc Phe	ccg Pro	gaa Glu 85	gga Gly	tat Tyr	aca Thr	tgg Trp	gag Glu 90	agg Arg	atc Ile	atg Met	aac Asn	ttt Phe 95	gaa Glu	2	88
gat Asp	ggt Gly	gct Ala	gtg Val 100	tgt Cys	act Thr	gtc Val	agc Ser	aat Asn 105	gat Asp	tcc Ser	agc Ser	atc Ile	caa Gln 110	ggc Gly	aac Asn	3	36
tgt Cys	ttc Phe	atc Ile 115	tac Tyr	cat His	gtc Val	aag Lys	ttt Phe 120	tct Ser	ggt Gly	ttg Leu	aac Asn	ttt Phe 125	cct Pro	ccc Pro	aat Asn	3	84
gga Gly	cct Pro 130	gtg Val	atg Met	cag Gln	aag Lys	aag Lys 135	aca Thr	cag Gln	ggc Gly	tgg Trp	gaa Glu 140	ccc Pro	aac Asn	act Thr	gag Glu	4	32
cgt Arg 145	ctc Leu	ttt Phe	gca Ala	cga Arg	gat Asp 150	gga Gly	atg Met	ctg Leu	ata Ile	gga Gly 155	aac Asn	aac Asn	ttt Phe	atg Met	gct Ala 160	4	80

#### 179/234

									179	/234						
ctg Leu	aag Lys	tta Leu	gaa Glu	gga Gly 165	ggc	ggt Gly	cac His	tat Tyr	ttg Leu 170	tgt Cys	gaa Glu	ttc Phe	aaa Lys	tcg Ser 175	act Thr	528
tac Tyr	aag Lys	gca Ala	aag Lys 180	aag Lys	act Thr	gtg Val	aag Lys	atg Met 185	cca Pro	ggg Gly	tat Tyr	cac His	tat Tyr 190	gtt Val	gac Asp	576
cgc Arg	aaa Lys	ctg Leu 195	gat Asp	gta Val	acc Thr	aat Asn	cac His 200	aac Asn	aag Lys	gat Asp	tac Tyr	act Thr 205	tcc Ser	gtt Val	gag Glu	624
cag Gln	tgt Cys 210	gaa Glu	att Ile	tcc Ser	att Ile	gca Ala 215	cgc Arg	aaa Lys	cct Pro	gtg Val	gtc Val 220	gcc Ala				663
<210	> 1	67														
<211	> 2	221					•									
<212	> I	PRT														
<213	> E	?avor	na de	ecuss	sata											
<400	> 1	67														
Met 1	Ser	Val	Ile	Ala 5	Thr	Gln	Val	Thr	Tyr 10	Lys	Val	Tyr	Met	Ser 15	Gly	
Thr	Val	Asn	Gly 20	His	Tyr	Phe	Glu	Val 25	Glu	Gly	Asp	Gly	Lys 30	Gly	Lys	
Pro	Tyr	Glu 35	Gly	Glu	Gln	Thr	Val 40	Arg	Leu	Thr	Val	Thr 45	Lys	Gly	Gly	
Pro	Leu 50	Pro	Phe		Trp	Asp 55	Ile	Leu	Ser	Pro	Gln 60	Tyr	Gln	Tyr	Glý	
Ser 65	Ile	Pro	Phe	Thr	Lys 70	Tyr	Pro	Glu	Asp	Ile 75		Asp	Tyr	Val	Lys 80	
Gln	Ser	Phe	Pro	Glu 85	Gly	Tyr	Thr	Trp	Glu 90	Arg	Ile	Met	Asn	Phe 95	Glu	
Asp	Gly	Ala	Val 100	Суѕ	Thr	Val	Ser	Asn 105	Asp	Ser	Ser	Ile	Gln 110	Gly	Asn	
Суѕ		Ile		His	Val	Lys	Phe	Ser	Gly	Leu	Asn	Phe	Pro	Pro	Asn	

115

120

125

#### 180/234

Gly Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu 130 135 140	
Arg Leu Phe Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala 145 150 155 160	
Leu Lys Leu Glu Gly Gly Gly His Tyr Leu Cys Glu Phe Lys Ser Thr 165 170 175	
Tyr Lys Ala Lys Lys Thr Val Lys Met Pro Gly Tyr His Tyr Val Asp 180 185 190	
Arg Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu 195 200 205	
Gln Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 210 215 220	
<210> 168	
<211> 660	
<212> DNA	
<213> Pavona decussata	
<220>	
<221> CDS .	
<222> (1)(660)	
<400> 168	
agt gtg atc gct aca caa atg acc tac aag gtt tat atg tca ggc acg Ser Val Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr	48
1 5 10 15	
gtc aat gga cac tac ttt gag gtt gaa ggc gat gga aaa gga aag cct Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys Pro	96
20 25 30	
tac gag ggg gag cag acg gta agg ctc act gtc aca aag ggc gga cct Tyr Glu Gly Glu Gln Thr Val Arg Leu Thr Val Thr Lys Gly Gly Pro 35 40 45	144
ctg cca ttt gct tgg gat att tta tca cca cag tat cag tac gga agc Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Tyr Gln Tyr Gly Ser 50 55 60	192

	w	<b>3</b> 02/0	70703	t												DCT/CD02/00020
	***	<i>y</i> 02/0	70792	•					181	/234						PCT/GB02/00928
ata Ile 65	cca Pro	ttc Phe	acc Thr	aag Lys	tac Tyr 70	cct Pro	gaa Glu	gạc Asp	atc Ile	cct Pro 75	gac Asp	tat Tyr	gta Val	aag Lys	cag Gln 80	240
tca Ser	ttc Phe	ccg Pro	gaa Glu	gga Gly 85	tat Tyr	aca Thr	tgg Trp	gag Glu	90 90	atc Ile	atg Met	aac Asn	ttt Phe	gaa Glu 95	gat Asp	288
ggt Gly	gct Ala	gtg Val	tgt Cys 100	act Thr	gtc Val	agc Ser	aat Asn	gat Asp 105	tcc Ser	agc Ser	atc Ile	caa Gln	ggc Gly 110	aac Asn	tgt Cys	336
ttc Phe	atc Ile	tac Tyr 115	cat His	gtc Val	aag Lys	ttc Phe	tct Ser 120	ggt Gly	ttg Leu	aac Asn	ttt Phe	cct Pro 125	ccc Pro	aat Asn	gga Gly	384
Pro	gtg Val 130	atg Met	cag Gln	aag Lys	aag Lys	aca Thr 135	cag Gln	Gly Ggc	tgg Trp	gaa Glu	ccc Pro 140	aac Asn	act Thr	gag Glu	cgt Arg	432
ctc Leu 145	ttt Phe	gca Ala	cga Arg	gat Asp	gga Gly 150	atg Met	ctg Leu	ata Ile	gga Gly	aac Asn 155	aac Asn	ttt Phe	atg Met	gct Ala	ctg Leu 160	480
aag Lys	tta Leu	gaa Glu	gga Gly	ggc Gly 165	ggt Gly	cac His	tat Tyr	ttg Leu	tgt Cys 170	gaa Glu	ttc Phe	aaa Lys	tcg Ser	act Thr 175	tac Tyr	528
aag Lys .	gca Ala	aag Lys	aag Lys 180	act Thr	gtg Val	aag Lys	atg Met	cca Pro 185	Gly	tat Tyr	cac His	tat Tyr	gtt Val 190	gac Asp	cgc Arg	576
aaa Lys	ctg Leu	gtt Val 195	gta Val	acc Thr	aat Asn	cac His	aac Asn 200	aag Lys	gat Asp	tac Tyr	act Thr	tcc Ser 205	gtt Val	gag Glu	cag Gln	624
tgt ( Cys (	gaa Glu 210	att Ile	tcc Ser	att Ile	gca Ala	cgc Arg 215	aaa Lys	cct Pro	gtg Val	gtc Val	gcc Ala 220					660
<210	> 1	69														
<211	> 2	20														
<212	> P	RT														
<213	> P	avon	a de	cuss	ata											
<400	> 1	69														
G 3	, ,	- 1		m ì												

Ser Val Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr 1 5 10 15

Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys Pro 20 25 30

#### 182/234

Tyr Glu Gly Glu Gln Thr Val Arg Leu Thr Val Thr Lys Gly Gly Pro 35 40 45

Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Tyr Gln Tyr Gly Ser 50 55 60

Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys Gln 65 70 75 80

Ser Phe Pro Glu Gly Tyr Thr Trp Glu Gly Ile Met Asn Phe Glu Asp 85 90 95

Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys 100 105 110

Phe Ile Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly 115 120 125

Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu Arg 130 135 140

Leu Phe Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu 145 150 155 160

Lys Leu Glu Gly Gly His Tyr Leu Cys Glu Phe Lys Ser Thr Tyr 165 170 175

Lys Ala Lys Lys Thr Val Lys Met Pro Gly Tyr His Tyr Val Asp Arg
180 185 190

Lys Leu Val Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu Gln
195 200 205

Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 210 215 220

<210> 170

<211> 663

<212> DNA

<213> Montipora sp.

<220>

183/234

<221> CDS

<222> (1)..(663)

atg	agt	170 gtg	atc	gct	aca	caa	atg	acc	tac	aag	gtt	tat	atg	tca	ggc	48
1				5					10			Tyr		15		
acg Thr	gtc Val	aat Asn	gga Gly 20	cac His	tac Tyr	ttt Phe	gag Glu	gtt Val 25	gaa Glu	ggc	gat Asp	gga Gly	aaa Lys 30	gga Gly	aag Lys	96
cct Pro	tac Tyr	gaa Glu 35	GJλ āāā	gag Glu	cag Gln	acg Thr	gta Val 40	agg Arg	ctc Leu	act Thr	gtc Val	aca Thr 45	aag Lys	ggċ Gly	gga Gly	144
cct Pro	ctg Leu 50	cca Pro	ttt Phe	gct Ala	tgg Trp	gat Asp 55	att Ile	tta Leu	tca Ser	cca Pro	cag Gln 60	tat Tyr	cag Gln	tac Tyr	gga Gly	192
Ser 65	Ile	Pro	Phe	Thr	Lys 70	Tyr	Pro	Glu	Asp	Ile 75	Pro	gac Asp	Tyr	Val	Lys 80	240
Gln	Ser	Phe	Pro	Glu 85	Gly	Tyr	Thr	Trp	Glu 90	Arg	Ile	atg Met	Asn	Phe 95	Glu	288
Asp	Gly	Ala	Val 100	Cys	Ala	Val	Ser	Asn 105	Asp	Ser	Ser	atc Ile	Gln 110	Gly	Asn	336
Суѕ	Phe	11e 115	Tyr	His	Val	Lys	Phe 120	Ser	Gly	Leu	Asn	ttt Phe 125	Pro	Pro	Asn	384
Gly	Pro 130	Val	Met	Gln	Lys	Lуs 135	Thr	Gln	Gly	Trp	Glu 140	ccc Pro	Asn	Thr	Glu	432
Arg 145	Leu	Phe	Ala	Arg	Asp 150	Gly	Met	Leu	Ile	Gly 155	Asn	aac Asn	Phe	Met	Ala 160	480
Leu	Lys	Leu	Glu	Gly 165	Gly	Gly	His	Tyr	Leu 170	Суѕ	Glu	ttc Phe	Lys	Ser 175	Thr	528
Tyr	Lys	Ala	Lys 180	Lys	Pro	Val	Lys	Met 185	Pro	Gly	Tyr	cac His	Tyr 190	Val	Āsp	576
cgc Arg	aaa Lys	ctg Leu 195	gat Asp	gta Val	acc Thr	Asn	cac His 200	aac Asn	aag Lys	gat Asp	tac Tyr	act Thr 205	tcc Ser	gtt Val	gag Glu	624

#### 184/234

Cag tgt gaa att tcc att gca cgc aaa cct gtg gtc gcc
Gln Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala
210 215 220

663

<210> 171

<211> 221

<212> PRT

<213> Montipora sp.

<400> 171

Met Ser Val Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Ser Gly
1 5 10 15

Thr Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys 20 25 30

Pro Tyr Glu Gly Glu Gln Thr Val Arg Leu Thr Val Thr Lys Gly Gly 35 40 45

Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Tyr Gln Tyr Gly 50 55 60

Ser Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys 65 70 75 80

Gln Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ile Met Asn Phe Glu 85 90 95

Asp Gly Ala Val Cys Ala Val Ser Asn Asp Ser Ser Ile Gln Gly Asn 100 105 110

Cys Phe Ile Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn 115 120 125

Gly Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu 130 135 140

Arg Leu Phe Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala 145 150 155 160

Leu Lys Leu Glu Gly Gly Gly His Tyr Leu Cys Glu Phe Lys Ser Thr 165 170 175

#### 185/234

Tyr	Lys	Ala	Lys	Lys	Pro	Val	Lys	Met	Pro	Gly	Tyr	His	Tyr	Val	Asp
			180					185					190		_

Arg Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu 195 200 205

Gln Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 210 215 220

<210> 172

<211> 663

<212> DNA

<213> Montipora sp.

<220>

<221> CDS

<222> (1)..(663)

<400> 1	172											
atg agt Met Ser 1												48
acg gtc Thr Val	aat gga Asn Gly 20	cac tac His Tyn	ttt g	ag gtt lu Val 25	gaa Glu	ggc Gly	gat Asp	gga Gly	aaa Lys 30	gga Gly	aag Lys	96
cct tac Pro Tyr				al Arg			-		_			144
cct ctg Pro Leu 50												192
agc ata Ser Ile 65				-	_					_	_	240
cag tca Gln Ser												288
gat ggt Asp Gly		Cys Thi										336

WO 02/070703	PCT/GB02/00928
	40.400

	WC	02/0	70703	3												PCT/GB02/009	)
									186	/234							
tgt Cys	ttc Phe	atc Ile 115	tac Tyr	cat His	gtc Val	aag Lys	ttc Phe 120	tct Ser	ggt Gly	ttg Leu	aac Asn	ttt Phe 125	cct Pro	ccc Pro	aat Asn	384	
gga Gly	cct Pro 130	gtg Val	atg Met	caa Gln	aaa Lys	aag Lys 135	aca Thr	caa Gln	ggc Gly	tgg Trp	gaa Glu 140	ccc Pro	aac Asn	act Thr	gag Glu	432	
cgt Arg 145	ctc Leu	ttt Phe	gca Ala	cga Arg	gat Asp 150	gga Gly	atg Met	ctg Leu	ata Ile	gga Gly 155	aac Asn	aac Asn	ttt Phe	atg Met	gct Ala 160	480	
ctg Leu	aag Lys	tta Leu	gaa Glu	gga Gly 165	ggc Gly	ggt Gly	cac His	tat Tyr	ttg Leu 170	tgt Cys	gaa Glu	ttc Phe	aaa Lys	tct Ser 175	act Thr	528	
tac Tyr	aag Lys	gca Ala	aag Lys 180	aag Lys	cct Pro	gtg Val	aag Lys	atg Met 185	cca Pro	GJ À GG À	tat Tyr	cac His	tat Tyr 190	gtt Val	gac Asp	576	
cgc Arg	aaa Lys	ctg Leu 195	gat Asp	gta Val	acc Thr	aat Asn	cac His 200	aac Asn	aag Lys	gat Asp	tac Tyr	act Thr 205	tcc Ser	gtt Val	GJ A āāā	624	
cag Gln	tgt Cys 210	gaa Glu	att Ile	tcc Ser	att Ile	gcc Ala 215	ccc Pro	aaa Lys	cct Pro	gtg Val	gtc Val 220	gcc Ala		•		663	
<210	)> 1	173															
<211	.> 2	221															
<212	2> I	PRT															
<213	3> N	1onti	pora	a sp.	•												
<400	)> ]	173															
Met 1	Ser	Val	Ser	Ala 5	Thr	Gln	Met	Thr	Tyr 10	Lys	Val	Tyr	Met	Ser 15	Gly		
Thr	Val	Asn	Gly	His	Tyr	Phe	Glu	Val	Glu	Gly	Asp	Gly	Lys	Gly	Lys		

Pro Tyr Glu Gly Glu Gln Thr Val Arg Leu Thr Val Thr Lys Gly Gly 

Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Tyr Gln Tyr Gly 50 55 60 

Ser Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Gly Tyr Val Lys 

#### 187/234

Gln Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ile Met Asn Phe Glu 85 90 95

Asp Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn 100 105 110

Cys Phe Ile Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn 115 120 125

Gly Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu 130 135 140

Arg Leu Phe Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala 145 150 155 160

Leu Lys Leu Glu Gly Gly Gly His Tyr Leu Cys Glu Phe Lys Ser Thr 165 170 175

Tyr Lys Ala Lys Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Asp 180 185 190

Arg Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Gly 195 200 205

Gln Cys Glu Ile Ser Ile Ala Pro Lys Pro Val Val Ala 210 215 220

<210> 174

<211> 660

<212> DNA

<213> Montipora sp.

<220>

<221> CDS

<222> (1)..(660)

<400> 174

agt gtg atc gct aca caa atg acc tac aag gtt tat atg tca ggc acg Ser Val Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr 1 5 10 15

WO 02/070703	PCT/GB02/00928
WO 02/0/0/03	PCT/GB02/00928

188/234 gtc aat gga cac tac ttt gag gtc gaa ggc gat gga aaa gga aag cct 96 Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys Pro tac gag ggg gag cag acg gta aag ctc act gtc acc aag ggc gga cct 144 Tyr Glu Gly Glu Gln Thr Val Lys Leu Thr Val Thr Lys Gly Gly Pro ctg cca ttt gct tgg gat att tta tca cca cag tgt cag tac gga agc 192 Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Cys Gln Tyr Gly Ser 55 ata cca ttc acc aag tac cct gaa gac atc cct gac tat gta aag cag 240 Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys Gln tca ttc ccg gag gga ttt aca tgg gag agg atc atg aac ttt gaa gat 288 Ser Phe Pro Glu Gly Phe Thr Trp Glu Arg Ile Met Asn Phe Glu Asp 85 90 ggt gca gtg tgt act gtc agc aat gat tcc agc atc caa ggc aac tgt 336 Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys tto acc tac cat gtc aag ttc tct ggt ttg aac ttt cct cct aat gga 384 Phe Thr Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly 120 cct gtg atg cag aag aca cag ggc tgg gaa ccc cac tct gag cgt 432 Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro His Ser Glu Arg 135 ctc ttt gca cgg ggt gga atg ctg ata gga aac aac ttt atg gct ctg 480 Leu Phe Ala Arg Gly Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu 150 aag tta gaa gga ggc ggt cac tat ttg tgt gaa ttc aaa act act tac 528 Lys Leu Glu Gly Gly His Tyr Leu Cys Glu Phe Lys Thr Thr Tyr 170 aag gca aag aag cct gtg aag atg cca ggg tat cat tat gtt gac cgc 576 Lys Ala Lys Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Asp Arg 185 aaa ctg gat gta acc aat cac aac aag gat tac act tcc gtt gag cag 624 Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu Gln

200

660

tgt gaa att tcc att gca cgc aaa cct gtg gtc gcc

Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala

<210> 175

210

<211> 220

<212> PRT

<213> Montipora sp.

195

<400> 175

Ser Val Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr

Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys Pro

Tyr Glu Gly Glu Gln Thr Val Lys Leu Thr Val Thr Lys Gly Gly Pro

Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Cys Gln Tyr Gly Ser 50 55 60

Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys Gln . 70

Ser Phe Pro Glu Gly Phe Thr Trp Glu Arg Ile Met Asn Phe Glu Asp 90

Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys 100 105 · 110

Phe Thr Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly 115 120

Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro His Ser Glu Arg

Leu Phe Ala Arg Gly Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu

Lys Leu Glu Gly Gly His Tyr Leu Cys Glu Phe Lys Thr Thr Tyr 165 170

Lys Ala Lys Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Asp Arg

Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu Gln

Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 215

#### 190/234

<210>	176														
<211>	660														
<212>	DNA									•					
<213>	Mont	ipor	a sp	•											
<220>															
<221>	CDS														
<222>	(1).	. (66	0)												
<400> agt gt	176	att	aca	caa	ato	300	tac	220	att	+-+	2+ <i>a</i>	+	~~~		40
Ser Va	l Ile	Val	Thr 5	Gln	Met	Thr	Tyr	Lys 10	Val	Tyr	Met	Ser	ggc Gly 15	Thr	48
gtc aa Val As	t gga n Gly	cac His	tac Tyr	ttt Phe	gag Glu	gtt Val	gaa Glu	ggc Gly	gat Asp	gga Glv	aaa Lvs	gga Glv	aag Lvs	cct Pro	96
		20					25	_	-	-	-	30	•		
tac ga Tyr Gl	a ggg u Gly 35	gag Glu	cag Gln	acg Thr	gta Val	agg Arg 40	ctc Leu	act Thr	gtc Val	aca Thr	aag Lys 45	ggc Gly	gga Gly	ccc Pro	144
ctg cc Leu Pr 50	o Phe	gct Ala	tgg Trp	gat Asp	att Ile 55	tta Leu	tca Ser	cca Pro	cag Gln	tat Tyr 60	cag Gln	tac Tyr	gga Gly	agc Ser	192
ata cc Ile Pr 65	a ttc o Phe	acc Thr	aag Lys	tac Tyr 70	cct Pro	gaa Glu	gac Asp	atc Ile	cct Pro 75	gac Asp	tat Tyr	gta Val	aag Lys	cag Gln 80	240
tca tt Ser Ph	c ccg e Pro	gaa Glu	gga Gly 85	tat Tyr	aca Thr	tgg Trp	gag Glu	agg Arg 90	atc Ile	atg Met	aac Asn	ttt Phe	gaa Glu 95	gat Asp	288
ggt gc Gly Al	a gtg a Val	tgt Cys 100	act Thr	gtc Val	agc Ser	aat Asn	gat Asp 105	tcc Ser	agc Ser	atc Ile	caa Gln	ggc Gly 110	aac Asn	tgt Cys	336
ttc at Phe Il	c tac e Tyr 115	His	gtc Val	aag Lys	ttc Phe	tct Ser 120	ggt Gly	ttg Leu	aac Asn	ttt Phe	cct Pro 125	ccc Pro	aat Asn	gga Gly	384
cct gt Pro Va 13	l Met	cag Gln	aag Lys	aag Lys	aca Thr 135	cag Gln	ggc Gly	tgg Trp	gaa Glu	ccc Pro 140	aac Asn	act Thr	gag Glu	cgt Arg	432
ctc tt Leu Ph 145	t gca e Ala	cga Arg	gat Asp	gga Gly 150	atg Met	ctg Leu	ata Ile	gga Gly	aac Asn 155	aac Asn	ttt Phe	atg Met	gct Ala	ctg Leu 160	480

# aag tta gaa gga ggc ggt cac tat ttg tgt gaa ttc aaa tct act tac 528 Lys Leu Glu Gly Gly Gly His Tyr Leu Cys Glu Phe Lys Ser Thr Tyr 175 aag gca aag aag cct gtg aag atg cca ggg tat cac tat gtt gac cgc Lys Ala Lys Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Asp Arg 180 aaa ctg gat gta acc aat cac aac aag gat tac acc tcc gtt gag cag Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu Gln 195 tgt gaa att tcc att gca cgc aaa cct gtg gtc gcc 660

<210> 177

<211> 220

<212> PRT

<213> Montipora sp.

<400> 177

Ser Val Ile Val Thr Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr 1 5 10 15

Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 210 215 220

Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys Pro 20 25 30

Tyr Glu Gly Glu Gln Thr Val Arg Leu Thr Val Thr Lys Gly Gly Pro 35 40 45

Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Tyr Gln Tyr Gly Ser 50 55

Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys Gln 65 70 75 80

Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ile Met Asn Phe Glu Asp 85 90 95

Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys 100 105 110

Phe Ile Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly 115 120 125

#### 192/234

Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu Arg 130 135 140

Leu Phe Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu 145 150 155 160

Lys Leu Glu Gly Gly Gly His Tyr Leu Cys Glu Phe Lys Ser Thr Tyr 165 170 175

Lys Ala Lys Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Asp Arg 180 185 190

Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu Gln 195 200 205

Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 210 215 220

<210> 178

<211> 701

<212> DNA

<213> Acanthastria sp.

<400> 178 teegttateg etaaacagat gacegettea aegttaagtt gacaacagga ageacgaegg 60 agactgcagt cccgtacgcg cgaacgggat acctgggatt tatcaagaga acagatttca 120 cgcagacaga tggagcccgg catgacgcgt tatttgtggt tggccctctt gaagaaacca 180 tgatattgcg tggtatgagg tatcacccgg tagatatcga gaacacagtg acgagatgtc 240 atcgatcaat ctgtgaaagt gcggtcttca cgatgacaaa cctacttgtg gtagcagtgg 300 agettgatge agatgaacge gaggeacttg aegtggttee getggtgaeg acateegtae 360 tgaatgaaca gcaacttgtc gtaggggtgg tggtagtggt tgaccctggc gtagtcccga 420 tcaattctcg cggagagaaa caacggatgc atctgaggga cgggttcctg ggggaccagt 480 tggatcctat ctacgtggcg tataatatgt agacacctca ctgcttagtt tcgtaattga 540 attgtgtcgt agtttttta aatgacaatt aatagacaag tttgaaattg actgtagcgc 600 taggtttagg tataaactag cgtttggtaa ggcaattatg acaggaacta ctgtcacgcg 660 tgacgcgaga ccgtcacttt acacgcaaac ctgtggtcgc c 701

# WO 02/070703 PCT/GB02/00928 ·

#### 193/234

<210>	179
<211>	701
<212>	DNA
<213>	Green Pocillopora

<400> 179 teegttateg etaaacagat gacegettea aegttaagtt gacaacagga ageacgacgg 60 agactgcagt cccgtacgcg cgaacgggat acctgggatt tatcaagaga acagatttca 120 cgcagacaga tggagcccgg catgacgcgt tatttgtggt tggccctctt gaagaaacca 180 tgatattgcg tggtatgagg tatcacccgg tagatatcga gaacacagtg acgagatgtc 240 atcgatcaat ctgtgaaagt gcggtcttca cgatgacaaa cctacttgtg gtagcagtgg 300 agettgatge agatgaacge gaggeacttg acgtggttcc gctggtgacg acatecgtae 360 tgaatgaaca gcaacttgtc gtaggggtgg tggtagtggt tgaccctggt gtagtcccga 420 tcaattctcg cggagagaaa caacggatgc atctgaggga cgggttcctg ggggaccagt 480 tggatcctat ctacgtggcg tataatatgt agacacctca ctgcttaatt ttcgtaattg 540 aattgtgtcg tagtttttt aaatgacaac taatagacag tttgaaattg actgtagcgc 600 taggtttagg tataaactag cgtttggtaa ggcaattatg acaggaatta ctgtcacgcg 660 tgacgcgaga ccgtcacttt acacgcaaac ctgtqgtcgc c 701

<210> 180

<211> 701

<212> DNA

<213> Green Pocillopora

<220>

<221> misc feature

<222> (634)...(634)

<223> n = any nucleotide

#### WO 02/070703 PCT/GB02/00928 194/234

<221> misc feature

<222> (640)..(640)

<223> n = any nucleotide

<400> 180		•				
tccgttatcg	ctaaacagat	gaccgcttca	acgttaagtt	gacaacagga	agcacgacgg	60
agactgcagt	cccgtacgcg	cgaacgggat	acctgggatt	tatcaagaga	acagatttca	120
cgcagacaga	tggagcccgg	catgacgcgt	tatttgtggt	tggccctctt	gaagaaacca	180
tgatattgcg	tggtatgagg	tatcacccgg	tagatatcga	gaacacagtg	acgagatgtc	240
atcgatcaat	ctgtgaaagt	gcggtcttca	cgatgacaaa	cctacttgtg	gtagcagtgg	300
agcttgatgc	agatgaacgc	gaggcacttg	acgtggttcc	gctggtgacg	acatecgtae	360
tgaatgaaca	gcaacttgtc	gtaggggtgg	tggtagtggt	tgaccctggc	gtagtcccga	420
tcaattctcg	cggagagaaa	caacggatgc	atctgaggga	cgggttcctg	ggggaccagt	480
tggatcctat	ctacgtggcg	tataatatgt	agacacctca	ctgcttagtt	tcgtaattga	540
attgtgtcgt	agtttttta	aatgacaatt	aatagacaag	tttgaaattg	actgtagcgc	600
taggtttagg	tataaactag	cgtttggtaa	ggcnattatn	acaggaacta	ctgtcacgcg	660
tgacgcgaga	ccgtcacttt	acacgcaaac	ctgtggtcgc	С		701
<210> 181						

<211> 701

<212> DNA

<213> Green Pocillopora

<400> 181 tccgttatcg ctaaacagat gaccgcttca ccgttaagtt gacaacagga agcacgacgg 60 agactgcagt cccgtacgcg cgaacgggat acctgggatt tatcaagaga acagatttca 120 cgcagacagg tggagcccgg catgacgcgt tatttgtggt tggccctctt gaagaaacca 180 tgatattgcg tggtatgagg tatcacccgg tagatatcga gaacacagtg acgagatgtc 240 atcgatcaat ctgtgaaagt gcggtcttca cgatgacaaa cctacttgtg gtagcagtgg 300 agettgatge agatgaacge gaggeacttg acgtggttee getggtgacg acateegtae 360 tgaatgaaca gcaacttgtc gtaggggtgg tggtagtggt tgaccctggt gtagtcccga 420

				195/234			
1	tcaattcto	cg cggagagaaa	caacggatgc	atctgaggga	cgggttcctg	ggggaccagt	480
1	tggatccta	it ctacgtggcg	tataatatgt	agacacctca	ctgcttagtt	tcgtaattga	540
ė	attgtgtc	gt agtttttta	aatgacaatt	aatagacaag	tttgaaattg	actgtagcgc	600
1	taggtttag	gg tataaactag	cgtttggtaa	ggcaattatg	acaggaatta	ctgtcacgcg	660
t	tgacgcga	ga ccgtcacttc	acacgcaaac	ctgtggtcgc	С		701
	<210> 18	32					
	<211> 70						
	<212> Di						
			/D				
•	<213> M	llepora sp.	(Hydrozoan)			•	•
•							
	<400> 18 tccgttate	32 og ctaaacagat	gaccgcttca	acgttaagtt	gacaacagga	agcacgacgg	60
ä	agactgcag	ıt cccgtacgcg	cgaacgggat	acctgggatt	tatcaagaga	acagatttca	120
ď	egcagacaç	ıg tggagcccgg	catgacgcgt	tatttgtggt	tggccctctt	gaagaaacca	180
t	tgatattgo	g tggtatgagg	tatcacccgg	tagatatcga	gaacacagtg	acgagatgtc	240
ć	atcgatcaa	ıt ctgtgaaagt	gcggtcttca	cgatgacaaa	cctacttgtg	gtagcagtgg	300
ć	agcttgato	jc agatgaacgc	gaggcacttg	acgtggttcc	gctggtgacg	acatccgtac	360
t	gtatgaad	a gcaacttgtc	gtaggggtgg	tggtagtggt	tgaccctggt	gtagtcccga	420
t	caattctc	g cggagagaaa	caacggatgc	atctgaggga	cgggttcctg	ggggaccagt	480
t	ggatecta	t ctacgtggcg	tataatatgt	agacacctca	ctgcttagtt	tcgtaattga	540
ā	attgtgtcg	ıt agtttttta	aatgacaatt	aatagacaag	tttgaaattg	actgtagcgc	600
t	aggtttag	g tataaactag	cgtttggtaa	ggcaattatg	acaggaatta	ctgtcacgcg	660
t	gacgcgag	a ccgtcacttc	acacgcaaac	ctgtggtcgc	С		701
	(210) 10	. 2					
	(210> 18						
	211> 70						
<	:212> DN	A.					

<220>

<221> CDS

<213> Pavona decussaca

#### 196/234

<222> (1)..(699)

-400	) \ 1	103														
	gtt		-		_	atg Met		_		_		-	tga		cag Gln 15	48
_	-	_		-	_	cag Gln			-		-		_			96
						att Ile										144
_	_		_		_	gcc Ala		-	_			-		tgc Cys		192
gta Val	tga			acc Thr		tag		tcg Ser	-		-	tga	-	gat Asp	-	240
						gtg Val										288
tgg Trp	tag					atg Met										336
		tgg Trp	_	_		ccg Pro		-	-					_	tag	384
_		tgg Trp				acc Thr							att Ile 130			432
	_				-	atc Ile	_	ggg Gly 140	-						-	480
			Ser	Thr	Trp	cgt Arg	Ile	Ile	Cys	Arg	His			-	tag	528
	cgt Arg		tga		gtg Val	tcg Ser	tag		ttt Phe 170	taa	_		att Ile		_	576
caa Gln		tga	aat Asn	tga	ctg Leu	tag	cgc Arg 180	tag	gtt Val	tag	gta Val	taa		agc Ser		624
tgg Trp	taa					agg Arg						tga		gag Glu		672

197/234

gtc act tta cac gca aac ctg tgg tcg cc Val Thr Leu His Ala Asn Leu Trp Ser 200 205

701

<210> 184

<211> 13

<212> PRT

<213> Pavona decussaca

<400> 184

<210> 185

<211> 46

<212> PRT

<213> Pavona decussaca

<400> 185

Gln Gln Glu Ala Arg Arg Arg Leu Gln Ser Arg Thr Arg Glu Arg Asp 1  $\phantom{000}$  5  $\phantom{000}$  10  $\phantom{000}$  15

Thr Trp Asp Leu Ser Arg Glu Gln Ile Ser Arg Arg Gln Met Glu Pro

Gly Met Thr Arg Tyr Leu Trp Leu Ala Leu Leu Lys Lys Pro 35 40 45

<210> 186

<211> 4

<212> PRT

<213> Pavona decussaca

<400> 186

Tyr Cys Val Val

<210> 187

<211> 4

<212> PRT

<213> Pavona decussaca

<400> 187

Gly Ile Thr Arg

<210> 188

<211> 5

<212> PRT

<213> Pavona decussaca

<400> 188

Ile Ser Arg Thr Gln 1 5

<210> 189

<211> 14

<212> PRT

<213> Pavona decussaca

<400> 189

<210> 190

<211> 5

<212> PRT

<213> Pavona decussaca

<400> 190

199/234

Gln Thr Tyr Leu Trp 1 5

<210> 191

<211> 17 .

<212> PRT

<213> Pavona decussaca

<400> 191

Gln Trp Ser Leu Met Gln Met Asn Ala Arg His Leu Thr Trp Phe Arg 1 5 10 15

Trp

<210> 192

<211> 4

<212> PRT

<213> Pavona decussaca

<400> 192

Arg His Pro Tyr

<210> 193

<211> 6

<212> PRT

<213> Pavona decussaca

<400> 193

Met Asn Ser Asn Leu Ser 1 5

<210> 194

<211> 5

200/234

<212> PRT

<213> Pavona decussaca

<400> 194

Trp Leu Thr Leu Ala 1 5

<210> 195

<211> 13

<212> PRT

<213> Pavona decussaca

<400> 195

Ser Arg Ser Ile Leu Ala Glu Arg Asn Asn Gly Cys Ile 1 5 10

<210> 196

<211> 23

<212> PRT

<213> Pavona decussaca

<400> 196

Ile Cys Arg His Leu Thr Ala

<210> 197

<211> 7

<212> PRT

<213> Pavona decussaca

<400> 197

Met Thr Ile Asn Arg Gln Val

<210> 198

<211> 4

<212> PRT

<213> Pavona decussaca

<400> 198

Thr Ser Val Trp

<210> 199

<211> 10

<212> PRT

<213> Pavona decussaca

<400> 199

<210> 200

<211> 12

<212> PRT

<213> Pavona decussaca

<400> 200

Arg Glu Thr Val Thr Leu His Ala Asn Leu Trp Ser 1  $\phantom{\bigg|}$  5  $\phantom{\bigg|}$  10

<210> 201

<211> 231

<212> PRT

<213> coral

<400> 201

Ser Val Ile Ala Lys Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr 1 5 10 15

Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys Pro 20 25 30

Tyr Glu Gly Glu Gln Thr Val Arg Leu Ala Val Thr Lys Gly Gly Pro 35 40 45

Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Cys Gln Tyr Gly Ser 50 55 60

Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys Gln 65 70 75 80

Ser Phe Pro Gly Arg Tyr Thr Trp Glu Arg Ile Met Asn Phe Glu Asp 85 90 95

Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys 100 105 110

Phe Ile Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly 115 120 125

Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu Arg 130 135 140

Leu Phe Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu 145 150 155 160

Lys Leu Glu Gly Gly His Tyr Leu Cys Glu Phe Lys Ser Thr Tyr 165 170 175

Lys Ala Arg Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Asp Arg 180 185 190

Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu Gln 195 200 205

Arg Glu Ile Ser Ile Ala Arg Lys Pro Leu Val Ala Cys Cys Phe Phe 210 215 220

Arg Val Lys Ser Arg His Lys 225 230

<210> 202

<211> 235

<212> PRT

<213> coral

<400> 202

#### 203/234

Ser Val Ile Ala Lys Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr 1 5 10 15

Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Leu Pro 20 25 30

Tyr Glu Gly Gly Gln Thr Val Arg Leu Ala Val Thr Lys Gly Gly Pro 35 45

Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Cys Gln Tyr Gly Ser 50 55 60

Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys Gln 65 70 75 80

Ser Phe Pro Gly Arg Tyr Thr Trp Glu Arg Ile Met Asn Phe Glu Asp 85 90 95

Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys 100 105 110

Phe Ile Tyr His Val Lys Arg Ser Gly Leu Asn Phe Pro Pro Asn Gly 115 120 125

Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu Arg 130 135 140

Leu Phe Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu 145 150 155 160

Lys Leu Glu Gly Gly His Tyr Leu Cys Glu Phe Lys Ser Thr Tyr 165 170 175

Lys Ala Lys Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Asp Arg 180 185 190

Lys Leu Asp Val Thr Asn His Asn Leu Asp Tyr Thr Ser Val Glu Gln 195 200 205

Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala Cys Arg Phe Phe 210 215 220

Arg Val Lys Ser Arg His Lys Tyr Ala Val Ala 225 230 235

<210> 203

<211> 49

<212> DNA

<213> oligonucleotide

<400> 203
tgagagaact agtctcgagc tctagaacaa gcttttttt tttttttt

	204/234	
<211>	41	
<212>	DNA	
<213>	oligonucleotide	
<400> cagggc	204 gcgc catgggatcc gttatcgcta aacagatgac c	41
<210>	205	
<211>	38	
<212>	DNA	
· <213>	oligonucleotide	
<400>		
gggtta	atta agctgcaggg cgaccacagg tttgcgtg	38
<210>	206	
<211>	18	
<212>	DNA	
<213>	oligonucleotide	
<400> cccgaa	206 aagt gccacctg	18
<210>		
<211>	19	
<212>	DNA	
<213>	oligonucleotide	
<400> gttctg	207 aggt cattactgg	. 19
<210>	208	
<211>	20	
<212>	DNA .	

PCT/GB02/00928

WO 02/070703

WO 02/070703		PCT/GB02/00928
	205/234	

<213> oligonucleotide

<400> 208 tcagggtact tggtgaatgg	20
<210> 209	
<211> 669	
<212> DNA	
<213> Acropora sp	
<400> 209 ggatccgtta tcgctaaaca gatgacctac aaggtttata tgtcaggcac ggtcaatgga	60
cactactttg aggtcgaagg cgatggaaaa ggaaagcctt acgaggggga gcagacggta	120
aagctcactg tcaccaaggg tggacctctg ccatttgctt gggatatttt atcaccacag	180
tcacagtacg gaagcatacc attcaccaag taccctgaag acatcccgga ctatgtaaag	240
cagtcattcc cggagggata tacatgggag aggatcatga actttgaaga tggtgcagtg	300
tgtactgtca gcaatgactc cagcatccaa ggcaactgtt tcatctacca tgtcaagttc	360
tctggtttga actttcctcc caatggacct gttatgcaga agaagacaca gggctgggaa	420
cccaacactg agcgtetett tgcacgagat ggaatgetga taggaaacaa etttatgget	480
ctgaagttag aaggaggtgg tcactatttg tgtgaattca aatctactta caaggcaaag	540
aagcctgtga agatgccagg gtatcactat gttgaccgca aactggatgt aaccaatcac	600
aacaaggatt acacttccgt tgagcagtgt gaaatttcca ttgcacgcaa acctgtggtc	660
gecetgeag	669
<210> 210 <211> 222	
<212> PRT	
<213> Acropora sp	
<400> 210	
Gly Ser Val Ile Ala Lys Gln Met Thr Tyr Lys Val Tyr Met Ser Gly 1 5 10 15	
Thr Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys 20 25 30	

## 206/234

:	
Pro Tyr Glu Gly Glu Gln Thr Val Lys Leu Thr Val Thr Lys Gly Gly 35 40	
Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Ser Gln Tyr Gly 50 60	
Ser Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys 75 80	
Gln Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ile Met Asn Phe Glu 85 90 95	
Asp Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn 100 105 110	
Cys Phe Ile Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn 115 120 125	
Gly Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu 130 135 140	
Arg Leu Phe Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Leu 145 150 155 160	
Lys Leu Glu Gly Gly Gly His Tyr Leu Cys Glu Phe Lys Ser Thr Tyr 165 170 175	
Lys Ala Lys Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Asp Arg 180 185 190	
Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu Gln 195 200 205	
Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala Leu Gln 210 215 220	
<210> 211	
<211> 669	
<212> DNA	
<213> Discosoma sp	
<400> 211	
	50
cactactttg aggtcgaagg cgatggaaaa ggaaagcctt acgaggggga gcagacggta 12	20
aggetgactg teaccaaggg eggacetetg ceatttgett gggatatttt ateaccaeag i 18	30
tcacagtacg gaagcatacc attcaccaag taccctgaag acatccctga ctatgtaaag 24	0
cagtcattcc cggagggata tacatgggag aggatcatga actttgaaga tggtgcagtg 30	)0
tgtactgtca gcaatgattc cagcatccaa ggcaactgtt tcatctacca tgtcaagttc 36	50
tctggtttga actttcctcc caatggacct gttatgcaga agaagacaca gggctgggaa 42	20

WO 02/	070703			207/234				PCT/GB02/00
cccaacactg	agcgtctc	tt agcac	gagat g	gaatgctga	taggaaa	caa c	tttatgg	et .480
ctgaagttag	aaggaggt	gg tcact	atttg t	gtgaattca	aatctac	tta c	aaggcaa	gg 540
aagcctgtga	agatgcca	gg gtato	actat g	tgaccgca	aactgga	tgt a	accaatca	ac 600
aacaaggatt	acacttcc	gt tgago	agcgt ga	aaatttcca	ttgcacg	caa a	cctgtgg	tc 660
gccctgcag								669
<210> 212								
<211> 222			•					
<212> PRT								
<213> Disc	cosoma sp	•						
					•			
<400> 212								
Gly Ser Val	l Ile Ala 5	Lys Glr	Met Th	r Tyr Lys 10	Val Tyr		Ser Gly 15	
Thr Val Ası	Gly His	Tyr Phe	Glu Va 25	-	Asp Gly	Lys 30	Gly Lys	
Pro Tyr Gla	ı Gly Glu	Gln Thi	Val Ar	g Leu Thr	Val Thr 45	Lys	Gly Gly	
Pro Leu Pro	o Phe Ala	Trp Asp 55	Ile Le	u Ser Pro	Gln Ser 60	Gln	Tyr Gly	
Ser Ile Pro	o Phe Thr	Lys Tyr	Pro Gl	u Asp Ile 75	Pro Asp	Tyr	Val Lys 80	

Gln Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ile Met Asn Phe Glu 90

Asp Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn 105

Cys Phe Ile Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn

Gly Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu

Arg Leu Leu Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Leu 155

Lys Leu Glu Gly Gly His Tyr Leu Cys Glu Phe Lys Ser Thr Tyr

Lys Ala Arg Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Asp Arg 185

Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu Gln 195 200

208/234

Arg Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala Leu Gln 210 215 220

<211> 669

<212> DNA

<213> Sinularia sp

<400> 213 ggatccgtta tcgctaaaca gatgacctac aaggtttata tgtcaggcac ggtcaatgga 60 cactactttg aggtcgaagg cgatggaaaa ggaaagcctt acgaggggga gcagacggta 120 aagctcactg tcaccaaggg tggacctctg ccatttgctt gggatatttt atcaccacag 180 tcacagtacg gaagcatacc attcaccaag taccctgaag acatcccgga ctatgtaaag 240 cagtcattcc cggaggggta tacatgggag aggatcatga actttgaaga tggtgcagtg 300 tgtactgtca gcaatgactc cagcatccaa ggcaactgtt tcatctacca tgtcaagttc 360 tetggtttga acttteette caatggaeet gttatgeaga agaagaeaca gggetgggaa 420 cccaacactg agcgtctctt tgcacgagat ggaatgctga taggaaacaa ctttatggct 480 ctgaagttag aaggaggtgg tcactatttg tgtgaattca aatctactta caaggcaaag 540 600 aageetgtga agatgeeagg gtateactat gttgaeegea aactggatgt aaccaateae aacaaggatt acactteegt tgageagtgt gaaattteea ttgeaegeaa acetttggte 660 669 gccctgcag

<210> 214

<211> 223

<212> PRT

<213> Sinularia sp

<400> 214

Gly Ser Val Ile Ala Lys Gln Met Thr Tyr Lys Val Tyr Met Ser Gly
1 5 10 15

Thr Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys 20 25 30

Pro Tyr Glu Gly Glu Gln Thr Val Lys Leu Thr Val Thr Lys Gly Gly 35 40 45

### 209/234

209/234	
Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Ser Gln Tyr Gly 50 55 60	
Ser Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys 65 70 75 80	
Gln Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ile Met Asn Phe Glu 85 90 95	
Asp Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn 100 105 110	
Cys Phe Ile Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Ser Asn 115 120 125	
Gly Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu 130 135 140	
Arg Leu Phe Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala 145 150 155 160	
Leu Lys Leu Glu Gly Gly Gly His Tyr Leu Cys Glu Phe Lys Ser Thr 165 170 175	
Tyr Lys Ala Lys Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Asp 180 185 190	
Arg Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu 195 200 205	
Gln Cys Glu Ile Ser Ile Ala Arg Lys Pro Leu Val Ala Leu Gln 210 215 220	
<210> 215	
<211> 669	
<212> DNA	
<213> Tubastrea sp	
<400> 215 ggatccgtta tcgctaaaca gatgacctac aaggtttata tgtcaggcac ggtcaatgg	ga

60 cactactttg aggtcgaagg cgatggaaaa ggaaagcctt acgaggggga gcagacggta 120 aagctcactg tcaccaaggg tggacctctg ccatttgctt gggatatttt atcaccacag 180 tcacagtacg gaagcatacc attcaccaag taccctgaag acatcccgga ctatgtaaag 240 cagtcattcc cggagggata tacatgggag aggatcatga actttgaaga tggtgcagtg 300 tgtactgtca gcaatgactc cagcatccaa ggcaactgtt tcatctacca tgtcaagttc 360 tctggtttga actttcctcc caatggacct gttatgcaga agaagacaca gggctgggaa 420 cccaacactg agcgtctctt tgcacgagat ggaatgctga taggaaacaa ctttatggct 480 ctgaagttag aaggaggtgg tcactatttg tgtgaattca aatctactta caaggcaaag 540

600

660 669

## 210/234

												•			
aago	ctg	tga a	agato	gccaç	gg gt	atca	actat	gtt	gaco	egca	aact	ggat	gt a	aacca	atcac
aaca	agga	att a	acact	tcc	jt to	gagca	agtgt	gaa	attt	cca	ttg	gcgc	caa a	accto	gtggtc
gccctgcag															
<210	<210> 216														
<211	.> 2	223													
<212	<212> PRT														
<213	<213> Tubastrea sp														
<400	)> :	216													
Gly 1	Ser	Val	Ile	Ala 5	Lys	Gln	Met	Thr	Tyr 10	Lys	Val	Tyr	Met	Ser 15	Gly
Thr	Val	Asn	Gly 20	His	Tyr	Phe	Glu	Val 25	Glu	Gly	Asp	Gly	Lуs. 30	Gly	Lys
Pro	Tyr	Glu 35	Gly	Glu	Gln	Thr	Val 40	Lys	Leu	Thr	Val	Thr 45	Lys	Gly	Gly
Pro	Leu 50	Pro	Phe	Ala	Trp	Asp 55	Ile	Leu	Ser	Pro	Gln 60	Ser	Gln	Tyr	Gly
Ser 65	Ile	Pro	Phe	Thr	Lys 70	Tyr	Pro	Glu	Asp	Ile 75	Pro	Asp	Tyr	Val	Lys 80
Gln	Ser	Phe	Pro	Glu 85	Gly	Tyr	Thr	Trp	Glu 90	Arg	Ile	Met	Asn	Phe 95	Glu
Asp	Gly	Ala	Val 100	Суз	Thr	Val	Ser	Asn 105	Asp	Ser	Ser	Ile	Gln 110	Gly	Asn
Cys	Phe	Ile 115	Туг	His	Val	Lys	Phe 120	Ser	Gly	Leu	Asn	Phe 125	Pro	Pro	Asn
Gly	Pro 130	Val	Met	Gln	Lys	Lys 135	Thr	Gln	Gly	Trp	Glu 140	Pro	Asn	Thr	Glu
Arg 145	Leu	Phe	Ala	Arg	Asp 150	Gly	Met	Leu	Ile	Gly 155	Asn	Asn	Phe	Met	Ala 160
Leu	Lys	Leu	Glu	Gly 165	Gly	Gly	His	Tyr	Leu 170	Cys	Glu	Phe	Lys	Ser 175	Thr
Tyr	Lys	Ala	Lys 180	Lys	Pro	Val	Lys	Met 185	Pro	Gly	Tyr	His	Tyr 190	Val	Asp
Arg	Lys	Leu 195	Asp	Val	Thr	Asn	His 200	Asn	Lys	Asp	Tyr	Thr 205	Ser	Val	Glu
Gln	Cys 210	Glu	I1e	Ser	Ile	Ala 215	Arg	Lys	Pro	Val	Val 220	Ala	Leu	Gln	· .

211/234

<210> 217 <211> 669 <212> DNA <213> Discosoma sp

<400> 217 ggatccgtta tcgctaaaca gatgacctac aaggtttata tgtcaggcac ggtcaatgga 60 cactactttg aggtcgaagg cgatggaaaa ggaaagcctt acgaggggga gcagacggta 120 aggetggetg teaceaaggg eggacetetg ceatttgett gggatatttt atcaceaeag 180 tgtcagtacg gaagcatacc attcaccaag taccctgaag acatccctga ctatgtaaag 240 cagtcattcc cggagggatt tacatgggag aggatcatga actttgaaga tggtgcagtg 300 tgtcctgtca gcaatgattc cagcatccaa ggcaactgtt tcatctacca tgtcaagttc 360 tctggtttga actttcctcc caatggacct gttatgcaga agaagacaca gggctgggaa 420 ccccactetg agcgtctctt tgcacgagac ggaatqctga taggaaacac ctttatggct 480 ctgaagttag aaggaggcgg tcactatttg tgtgaattca aaactactta caaggcaaag 540 aagcctgtga agatgccagg gtatcattat gttgaccgca aactggatgt aatcaatcac 600 aacaaggatt acacttccgt tgagcagtgt gaaatttcca ttgcacgcaa acctgtgqtc 660 gccctgcag 669

<210> 218

<211> 223

<212> PRT

<213> Discosoma sp

<400> 218

Gly Ser Val Ile Ala Lys Gln Met Thr Tyr Lys Val Tyr Met Ser Gly
1 5 10 15

Thr Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys 20 25 30

Pro Tyr Glu Gly Glu Gln Thr Val Arg Leu Ala Val Thr Lys Gly Gly 35 40 45

Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Cys Gln Tyr Gly 50 55 60

### 212/234

												•				
Ser 65	Ile	Pro	Phe		Lys 70	Tyr	Pro	Glu	Asp	Ile 75	Pro	Asp	Tyr	Val	Lys 80	
Gln	Ser	Phe	Pro	Glu 85	Gly	Phe	Thr	Trp	Glu 90	Arg	Ile	Met	Asn	Phe 95	Glu	
Asp	Gly	Ala	Val 100	Суѕ	Pro	Val	Ser	Asn 105	Asp	Ser	Ser	Ile	Gln 110	Gly	Asn	
Cys	Phe	Ile 115	Tyr	His	Val	Lys	Phe 120	Ser	Gly	Leu	Asn	Phe 125	Pro	Pro	Asn	
Gly	Pro 130	Val	Met	Gln	Lys	Lys 135	Thr	Gln	Gly	Trp	Glu 140	Pro	His	Ser	Glu	
Arg 145	Leu	Phe	Ala	Arg	Asp 150	Gly	Met	Leu	Ile	Gly 155	Asn	Thr	Phe	Met	Ala 160	
Leu	Lys	Leu	Glu	Gly 165	Gly	Gly	His	Tyr	Leu 170	Cys	Glu	Phe	Lys	Thr 175	Thr	
Tyr	Lys	Ala	Lys 180	Lys	Pro	Val	Lys	Met 185	Pro	Gly	Tyr	His	Tyr 190	Va1	Asp	
Arg	Lys	Leu 195	Asp	Val	Ile	Asn	His 200	Asn	Lys	Asp	Tyr	Thr 205	Ser	Val	Glu	
Gln	Cys 210	Glu	Ile	Ser	Ile	Ala 215	Arg	Lys	Pro	Val	Val 220	Ala	Leu	Gln		
<210	)>	219														
<211	L>	555														
<212	2>	DNA														
<213	3>	Sinu	lari	a sp					•							
<400	)>	219														

acggtaaggc tggctgtcac caagggcgga cctctgccat ttgcttggga tattttatca 60 ccacagtgtc agtacggaag cataccattc accaagtacc ttgaagacat ccctgactat 120 180 gtaaagcagt cattcccgga gggatttaca tgggagagga tcatgaactt tgaagatggt gcagtgtgta ctgtcagcaa tgattccagc atccaaggca actgtttcat ctaccatgtc 240 300 aagttctctg gtttgaactt tcctcccaat ggacctgtta tgcagaagaa gacacagggc 360 tgggaaccca acactgageg tetetttgca egagatggaa tgetgatagg aaacaacttt atggctctaa agttagaggg aggtggtcac tatttgtgtg aattcaaatc tacttacaag 420 480 gcaaagaagc ctgtgaagat gccagggtat cactatgttg accgcaaact ggatgtaacc 540 aatcacaaca aggattacac ttccgttgag cagtgtgaaa tttccattgc acgcaaacct ttggtcgccc tgcag 555

213/234

<210> 220

<211> 223

<212> PRT

<213> Sinularia sp

<400> 220

Gly Ser Val Ile Ala Lys Gln Met Thr Tyr Lys Val Tyr Met Ser Gly
1 5 10 15

Thr Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys 20 25 30

Pro Tyr Glu Gly Glu Gln Thr Val Arg Leu Ala Val Thr Lys Gly Gly 35 40 45

Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Cys Gln Tyr Gly 50 55 60

Ser Ile Pro Phe Thr Lys Tyr Leu Glu Asp Ile Pro Asp Tyr Val Lys 65 70 75 80

Gln Ser Phe Pro Glu Gly Phe Thr Trp Glu Arg Ile Met Asn Phe Glu 85 90 95

Asp Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn 100 105 110

Cys Phe Ile Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn 115 120 125

Gly Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu 130 135 140

Arg Leu Phe Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala 145 150 155 160

Leu Lys Leu Glu Gly Gly Gly His Tyr Leu Cys Glu Phe Lys Ser Thr 165 170 175

Tyr Lys Ala Lys Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Asp 180 185 190

Arg Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu 195 200 205

Gln Cys Glu Ile Ser Ile Ala Arg Lys Pro Leu Val Ala Leu Gln 210 215 220

<210> 221

<211> 669

<212> DNA

<213> Tubastrea sp

	221 tta	tcgctaaaca	gatgacctac	aaggtttata	tgtcaggcac	ggtcaatgga	60
cactact	ttg	aggtcgaagg	cgatggaaaa	ggaaagcctt	acgaggggga	gcagacggta	120
aggctgg	ctg	tcaccaaggg	cggacctctg	ccatttgctt	gggatatttt	atcaccacag	180
tgtcagt	acg	gaagcatacc	attcaccaag	taccctgaag	acatccctga	ctatgtaaag	240
cggtcat	tcc	cggagggatt	tacatgggag	aggatcatga	actttgaaga	tggtgcagtg	300
tgtactg	tca	gcaatgattc	cagcatccaa	ggcaactgtt	tcatctacca	tgtcaagttc	360
tctggtt	tga	actttcctcc	caatggacct	gttatgcaga	agaagacaca	gggctgggaa	420
ccccact	ctg	agcgtctctt	tgcacgagac	ggaatgctga	taggaaacaa	ctttatggct	480
ctgaagt	tag	aaggaggcgg	tcactatttg	tgtgaattca	aaactactta	caaggcaaag	540
aagcctg	tga	agatgccagg	gtatcattat	gttgaccgca	aactggatgt	aatcaatcac	600
aacaagg	att	acacttccgt	tgagcagtgt	gaaatttcca	ttgcacgcaa	acctgtggtc	660
gccctgc	ag						669
<210>	222						
.044.							

<211> 223

<212> PRT

<213> Tubastrea sp

<400> 222

Gly Ser Val Ile Ala Lys Gln Met Thr Tyr Lys Val Tyr Met Ser Gly 1 5 10 15

Thr Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys 20 25 30

Pro Tyr Glu Glu Gln Thr Val Arg Leu Ala Val Thr Lys Gly Gly 35 40 45

Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Cys Gln Tyr Gly 50 55 60

Ser Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys 65 70 75 80

Arg Ser Phe Pro Glu Gly Phe Thr Trp Glu Arg Ile Met Asn Phe Glu 85 90 95

#### 215/234

Asp Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn 100 105 110

Cys Phe Ile Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn 115 120 125

Gly Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro His Ser Glu 130 135 140

Arg Leu Phe Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala 145 150 155 160

Leu Lys Leu Glu Gly Gly Gly His Tyr Leu Cys Glu Phe Lys Thr Thr 165 170 175

Tyr Lys Ala Lys Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Asp 180 185 190

Arg Lys Leu Asp Val Ile Asn His Asn Lys Asp Tyr Thr Ser Val Glu 195 200 205

Gln Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala Leu Gln 210 215 220

<210> 223

<211> 46

<212> DNA

<213> oligonucleotide

<400> 223

cagggcgcgc caaggagata taacaatggc ttcctcagtt ctttcc

46

<210> 224

<211> 33

<212> DNA

<213> oligonucleotide

<400> 224

cactggatcc gcattgcact cttccgccgt tqc

33

<210> 225

<211> 45

<212> DNA

<213> oligonucleotide

<400> gcatgg	225 cgcg ccaaggagat ataacaatga agactaatct ttttc	45
<210>		
<211>	34	
<212>	DNA	
<213>	oligonucleotide	
<400> gcatgg	226 atcc gaattcggcc gaggataatg atag	34
<210>	227	
<211>	45	
<212>	DNA	
<213>	oligonucleotide	
<400> gcatgg	227 cgcg ccaaggagat ataacaatga agactaatct ttttc	45
<210>	228	
<211>	34	
<212>	DNA	
<213>	oligonucleotide	
	228 atcc gaattcggcc gaggataatg atag	34
<210>	229	
<211>	46	
<212>	DNA	
<213>	oligonucleotide	
<400> gatctt	229 aatt aaagctcatc atgctgcagg gcgaccacag gtttgc	46

<210>	230	
<211>	42	
<212>	DNA	
<213>	oligonucleotide	
<400> gcatcto	230 gcag gtcgccacca gtaaaggaga agaacttttc ac	42
<210>	231	
<211>	39	
<212>	DNA	
<213>	oligonucleotide	
<400> ctgatta	231 aatt aattatttgt atagttcatc catgccatg	39
<210>	232	
<211>	55	
<212>	DNA	
<213>	oligonucleotide	
<400> cagggco	232 gcgc caaggagata taacaatggg atccgttatc gctaaacaga tgacc	55
<210>	233	
<211>	45	
<212>	DNA	
<213>	oligonucleotide	
<400> ggctcta	233 agaa aggagatata caatgtccgt tatcgctaaa cagat	45

<210> 234

<211> 45

218/234

<212> DNA

<213> oligonucleotide

<400> 234
ggctctagaa aggagatata caatgtccgt tatcqctaaa caqat

45

<210> 235

<211> 50

<212> DNA

<213> oligonucleotide

<400> 235

ggcaagcttt cagtggtggt ggtggtggtg ggcgaccaca ggtttgcgtg

50

<210> 236

<211> 221

<212> PRT

<213> coral

<400> 236

Met Ser Val Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Ser Gly
1 5 10 15

Thr Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys 20 25 30

Pro Tyr Glu Gly Glu Gln Thr Val Lys Leu Thr Val Thr Lys Gly Gly 35 40 45

Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Cys Gln Tyr Gly 50 55 60

Ser Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys 65 70 75 80

Gln Ser Phe Pro Glu Gly Phe Thr Trp Glu Arg Ile Met Asn Phe Glu 85 90 95

Asp Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn 100 105 110

Cys Phe Thr Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn 115 120 125

#### 219/234

Gly Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Ser Ser Glu 130 135 140

His Leu Phe Ala Arg Gly Gly Met Leu Ile Gly Asn Asn His Met Ala 145 150 155 160

Leu Lys Leu Glu Gly Gly His Tyr Leu Cys Glu Phe Lys Thr Thr 165 170 175

Tyr Lys Ala Lys Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Asp 180 185 190

Arg Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu 195 200 205

Gln Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 210 215 220

<210> 237

<211> 221

<212> PRT

<213> coral

<400> 237

Met Ser Val Ile Ala Thr Gln Met Thr Tyr Lys Val Tyr Met Ser Gly
1 5 10 15

Thr Val Asn Gly His Tyr Phe Glu Val Gln Gly Asp Gly Lys Gly Lys 20 25 30

Pro Tyr Glu Glu Glu Gln Thr Val Lys Leu Thr Val Thr Lys Gly Gly 35 40 45

Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Cys Gln Tyr Gly 50 55

Ser Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys 65 70 75 80

Gln Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ile Met Asn Phe Glu 85 90 95

Asp Gly Ala Val Cys Thr Val Ser Asp Ser Ser Ile Gln Gly Asn 100 105 110

Cys Phe Ile Tyr Asn Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn 115 120 125

Gly Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu 130 135 140

Arg Leu Phe Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala 145 150 155 160

#### 220/234

Leu Lys Leu Glu Gly Gly Gly His Tyr Leu Cys Glu Phe Lys Ser Thr 165 170 175

Tyr Lys Ala Lys Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Asp 180 185 190

Arg Lys Leu Asp Val Thr Asn His Asn Ile Asp Tyr Thr Ser Val Glu 195 200 205

Gln Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala 210 215 220

<210> 238

<211> 226

<212> PRT

<213> coral

<400> 238

Ser Val Ile Ala Lys Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr 1 5 10 15

Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys Pro 20 25 30

Tyr Glu Gly Glu Gln Thr Val Arg Leu Thr Val Thr Lys Gly Gly Pro 35 40 45

Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Ser Gln Tyr Gly Ser 50 55 60

Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys Gln 65 70 75 80

Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ile Met Asn Phe Glu Asp 85 90 95

Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys 100 105 110

Phe Ile Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly 115 120 125

Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu Arg 130 135 140

Leu Phe Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu 145 150 155 160

Lys Leu Glu Gly Gly His Tyr Leu Cys Glu Phe Lys Ser Thr Tyr 165 170 175

Lys Ala Lys Lys Pro Val Arg Met Pro Gly Tyr His Tyr Val Asp Arg 180 185 190

#### 221/234

Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu.Gln 195 200 205

Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala Trp Cys Phe Phe 210 215 220

Arg Val 225

<210> 239

<211> 220

<212> PRT

<213> coral

<400> 239

Ser Val Ile Ala Lys Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr 1 5 10 15

Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys Pro 20 25 30

Tyr Glu Gly Glu Gln Thr Val Arg Leu Ala Val Thr Lys Gly Gly Pro 35 40 45

Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Cys Gln Tyr Gly Ser 50 55 60

Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys Gln
65 70 75 80

Ser Phe Pro Gly Arg Tyr Thr Trp Glu Arg Ile Met Asn Phe Glu Asp 85 90 95

Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys 100 105 110

Phe Ile Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly 115 120 125

Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu Arg 130 135 140

Leu Phe Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu 145 150 155 160

Lys Leu Glu Gly Gly His Tyr Leu Cys Glu Phe Lys Ser Thr Tyr 165 170 175

Lys Ala Arg Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Asp Arg 180 185 190

Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu Gln 195 200 205

#### 222/234

Arg Glu Ile Ser Ile Ala Arg Lys Pro Leu Val Ala 210 215 220

<210> 240

<211> 230

<212> PRT

<213> coral

<400> 240

Met Ser Cys Ser Lys Asn Val Ile Lys Glu Phe Met Arg Phe Lys Val 1 5 10 15

Arg Met Glu Gly Thr Val Asn Gly His Glu Phe Glu Ile Lys Gly Glu 20 25 30

Gly Glu Gly Arg Pro Tyr Glu Gly His Cys Ser Val Lys Leu Met Val 35 40 45

Thr Lys Gly Gly Pro Leu Pro Phe Ala Phe Asp Ile Leu Ser Pro Gln 50 55 60

Phe Gln Tyr Gly Ser Lys Val Tyr Val Lys His Pro Ala Asp Ile Pro 65 70 75 80

Asp Tyr Lys Lys Leu Ser Phe Pro Glu Gly Phe Lys Trp Glu Arg Val 85 90 95

Met Asn Phe Glu Asp Gly Gly Val Val Thr Val Ser Gln Asp Ser Ser 100 105 110

Leu Lys Asp Gly Cys Phe Ile Tyr Glu Val Lys Phe Ile Gly Val Asn 115 120 125

Phe Pro Ser Asp Gly Pro Val Met Gln Arg Arg Thr Arg Gly Trp Glu 130 135 140

Ala Ser Ser Glu Arg Leu Tyr Pro Arg Asp Gly Val Leu Lys Gly Asp 145 150 155 160

Ile His Met Ala Leu Arg Leu Glu Gly Gly His Tyr Leu Val Glu 165 170 175

Phe Lys Ser Ile Tyr Met Val Lys Lys Pro Ser Val Gln Leu Pro Gly 180 185 190

Tyr Tyr Tyr Val Asp Ser Lys Leu Asp Met Thr Ser His Asn Glu Asp 195 200 205

Tyr Thr Val Val Glu Gln Tyr Glu Lys Thr Gln Gly Arg His His Pro 210 215 220

Phe Ile Lys Pro Leu Gln 225 230

<210> 241

<211> 225

<212> PRT

<213> coral

<400> 241

Met Arg Ser Ser Lys Asn Val Ile Lys Glu Phe Met Arg Phe Lys Val 1 5 10 15

Arg Met Glu Gly Thr Val Asn Gly His Glu Phe Glu Ile Glu Gly Glu 20 25 30

Gly Glu Gly Arg Pro Tyr Glu Gly His Asn Thr Val Lys Leu Lys Val 35 40 45

Thr Lys Gly Gly Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln 50 55 60

Phe Gln Tyr Gly Asn Lys Val Tyr Val Lys His Pro Ala Asp Ile Pro 65 70 75 80

Asp Tyr Lys Lys Leu Ser Phe Pro Glu Gly Phe Lys Trp Glu Arg Trp 85 90 95

Met Asn Phe Glu Asp Gly Gly Val Val Thr Val Thr Gln Asp Ser Ser 100 105 110

Leu Gln Asp Gly Cys Phe Ile Tyr Lys Val Lys Phe Ile Gly Val Asn 115 120 125

Phe Pro Ser Asp Gly Pro Val Met Gln Lys Lys Thr Met Gly Trp Glu 130 135 140

Ala Ser Thr Lys Arg Leu Tyr Pro Arg Asp Gly Val Leu Lys Gly Glu 145 150 155 160

Ile His Lys Ala Leu Lys Leu Lys Asp Gly Gly His Tyr Leu Val Glu 165 170 175

Phe Lys Ser Ile Tyr Met Ala Lys Lys Pro Val Gln Leu Pro Gly Tyr 180 185 190

Tyr Tyr Val Asp Ser Lys Leu Asp Ile Thr Ser His Asn Glu Asp Tyr 195 200 205

Thr Ile Val Glu Gln Tyr Glu Arg Thr Glu Gly Arg His His Leu Phe 210 215 220

Leu

225

<210> 242

<211> 230

<212> PRT

<213> coral

<400> 242

Met Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu Val 1 5 10 15

Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly Glu 20 25 30

Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile Cys 35 40 45

Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr Phe 50 55 60

Ser Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys Arg 65 70 75 80

His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu Arg . 85 90 95

Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu Val 100 105 110

Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly Ile 115 120 125

Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr Asn 130 135 140

Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn Gly 145 150 155 160

Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser Val 165 170 175

Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly Pro 180 185 190

Val Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu Ser 195 200 205

Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe Val 210 215 220

Thr Ala Ala Gly Ile Thr

<210> 243

<211> 818

<212> DNA

<213> Aequorea victoria

<400> 243						
ggatccaagg	agatataaca	atgaagacta	atcttttct	ctttctcatc	ttttcacttc	60
tcctatcatt	atcctcggcc	gaattcagta	aaggagaaga	acttttcact	ggagttgtcc	120
caattcttgt	tgaattagat	ggtgatgtta	atgggcacaa	attttctgtc	agtggagagg	180
gtgaaggtga	tgcaacatac	ggaaaactta	cccttaaatt	tatttgcact	actggaaaac	240
tacctgttcc	atggccaaca	cttgtcacta	ctttctctta	tggtgttcaa	tgcttttcaa	300
gatacccaga	tcatatgaag	cggcacgact	tcttcaagag	cgccatgcct	gagggatacg	360
tgcaggagag	gaccatcttc	ttcaaggacg	acgggaacta	caagacacgt	gctgaagtca	420
agtttgaggg	agacaccctc	gtcaacagga	tcgagcttaa	gggaatcgat	ttcaaggagg	480
acggaaacat	cctcggccac	aagttggaat	acaactacaa	ctcccacaac	gtatacatca	540
tggcagacaa	acaaaagaat	ggaatcaaag	ttaacttcaa	aattagacac	aacattgaag	600
atggaagcgt	tcaactagca	gaccattatc	aacaaaatac	tccaattggc	gatggccctg	. 660
tccttttacc	agacaaccat	tacctgtcca	cacaatctgc	cctttcgaaa	gatcccaacg	720
aaaagagaga	ccacatggtc	cttcttgagt	ttgtaacagc	tgctgggatt	acacatggca	780
tggatgaact	atacaaacat	gatgagcttt	aagagctc			818

<210> 244

<211> 263

<212> PRT

<213> Aequorea victoria

### <400> 244

Met Lys Thr Asn Leu Phe Leu Phe Leu Ile Phe Ser Leu Leu Ser

Leu Ser Ser Ala Glu Phe Ser Lys Gly Glu Glu Leu Phe Thr Gly Val

Val Pro Ile Leu Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe

Ser Val Ser Gly Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr

Leu Lys Phe Ile Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr

Leu Val Thr Thr Phe Ser Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro

Asp His Met Lys Arg His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly 100 105 110

Tyr Val Gln Glu Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys 115 120 125

Thr Arg Ala Glu Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile 130 135 140

Glu Leu Lys Gly Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His 145 150 155 160

Lys Leu Glu Tyr Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp 165 170 175

Lys Gln Lys Asn Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile 180 185 190

Glu Asp Gly Ser Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro 195 200 205

Ile Gly Asp Gly Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr 210 215 220

Gln Ser Ala Leu Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val 225 230 235 240

Leu Leu Glu Phe Val Thr Ala Ala Gly Ile Thr His Gly Met Asp Glu 245 250 255

Leu Tyr Lys His Asp Glu Leu 260

<210> 245

<211> 235

<212> PRT

<213> Acropora aspera

<400> 245

Ser Val Ile Ala Lys Gln Met Thr Tyr Lys Val Tyr Met Ser Gly Thr 1 5 10 15

Val Asn Gly His Tyr Phe Glu Val Glu Gly Asp Gly Lys Gly Lys Pro 20 25 30

Tyr Glu Gly Glu Gln Thr Val Arg Leu Ala Val Thr Lys Gly Gly Pro 35 40 45

Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Cys Gln Tyr Gly Ser 50 55 60

Ile Pro Phe Thr Lys Tyr Pro Glu Asp Ile Pro Asp Tyr Val Lys Gln 65 70 75 80

#### 227/234

Ser Phe Pro Gly Arg Tyr Thr Trp Glu Arg Ile Met Asn Phe Glu Asp 85 90 95

Gly Ala Val Cys Thr Val Ser Asn Asp Ser Ser Ile Gln Gly Asn Cys
100 105 110

Phe Ile Tyr His Val Lys Phe Ser Gly Leu Asn Phe Pro Pro Asn Gly
115 120 125

Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Asn Thr Glu Arg 130 135 140

Leu Phe Ala Arg Asp Gly Met Leu Ile Gly Asn Asn Phe Met Ala Leu 145 155 160

Lys Leu Glu Gly Gly His Tyr Leu Cys Glu Phe Lys Ser Thr Tyr 165 170 175

Lys Ala Lys Lys Pro Val Lys Met Pro Gly Tyr His Tyr Val Asp Arg 180 185 190

Lys Leu Asp Val Thr Asn His Asn Lys Asp Tyr Thr Ser Val Glu Gln . 195 200 205

Cys Glu Ile Ser Ile Ala Arg Lys Pro Val Val Ala Cys Arg Phe Phe 210 215 220

Arg Val Lys Ser Arg His Lys Val Ala Val Ala 225 230 235

<210> 246

<211> 232

<212> PRT

<213> Acropora aspera

<400> 246

Met Ala Ser Phe Leu Lys Lys Thr Met Pro Phe Lys Thr Thr Ile Glu 1 5 10 15

Gly Thr Val Asn Gly His Tyr Phe Lys Cys Thr Gly Lys Gly Glu Gly
20 25 30

Asn Pro Phe Glu Gly Thr Gln Glu Met Lys Ile Glu Val Ile Glu Gly 35 40 45

Gly Pro Leu Pro Phe Ala Phe His Ile Leu Ser Thr Ser Cys Met Tyr 50 55 60

Gly Ser Lys Thr Phe Ile Lys Tyr Val Ser Gly Ile Pro Asp Tyr Phe 65 70 75 80

Lys Gln Ser Phe Pro Glu Gly Phe Thr Trp Glu Arg Thr Thr Tyr 85 90 95

### 228/234

Glu Asp Gly Gly Phe Leu Thr Ala His Gln Asp Thr Ser Leu Asp Gly 105 Asp Cys Leu Val Tyr Lys Val Lys Ile Leu Gly Asn Asn Phe Pro Ala 120 Asp Gly Pro Val Met Gln Asn Lys Ala Gly Arg Trp Glu Pro Ala Thr Glu Ile Val Tyr Glu Val Asp Gly Val Leu Arg Gly Gln Ser Leu Met Ala Leu Lys Cys Pro Gly Gly Arg His Leu Thr Cys His Leu His Thr Thr Tyr Arg Ser Lys Lys Pro Ala Ser Ala Leu Lys Met Pro Gly Phe His Phe Glu Asp His Arg Ile Glu Ile Met Glu Glu Val Glu Lys Gly 200 Lys Cys Tyr Lys Gln Tyr Glu Ala Ala Val Gly Arg Tyr Cys Asp Ala Ala Pro Ser Lys Leu Gly His Asn 230 <210> 247 <211> 51 <212> DNA <213> oligonucleotide cgcgccaagg agatataaca atgagaggat cgcatcacca tcaccatcac q 51 <210> 248 <211> 51 <212> DNA <213> oligonucleotide <400> 248 gatccgtgat ggtgatggtg atgcgatcct ctcattgtta tatctccttg g 51 <210> 249 <211> 47

<212> DNA

	229/234	
<213>	oligonucleotide	
<400> ctgatta	249 aatt aaageteate atgtttgtat agtteateea tgeeatg	47
<210>	250	
<211>	34	
<212>	DNA ·	
<213>	oligonucleotide	
<400> gtgtgt	250 actg tcagccagga ttccagcatc caag	34
<210>	251	
<211>	32	
<212>	DNA	
<213>	oligonucleotide	
<400> ctgtca	251 gcaa tgatatcagc atccaaggca ac	32
<210>	252	
<211>	4 4	
<212>	DNA	
<213>	oligonucleotide	
<400> ggatcca	252 atcg ccaccatgtc taaaggtgaa gaattattca ctgg	44
<210>	253	
<211>	34	
<212>	DNA	
<213>	oligonucleotide	
<400>	253	

PCT/GB02/00928

WO 02/070703

V	VO 02/070703 230/234	PCT/GB02/00928
cagctgt	tat ttgtacaatt catccatacc atgg	34
<210>	254	
<211>	41	
<212>	DNA	
<213>	oligonucleotide	
<400> cgggato	254 ccat cgccaccatg aggtcttcca agaatgttat c	41
<210>	255 .	
<211>	31	
<212>	DNA	
<213>	oligonucleotide	·
<400> gaggato	255 cege ggeegetaaa ggaacagatg g	31
<210>	256	•
<211>	38	
<212>	DNA	
<213>	oligonucleotide	
<400> gaagato	256 ctaa aacaatgagt gtgatcgcta cacaaatg	38
<210>	257	
<211>	35	
<212>	DNA	
<213>	oligonucleotide	
<400> tatcaa	257 atcg ceggegteag gegaceaeag gtttg	35
<210>	258	

	231/234			
<211>	30			
<212>	DNA			
<213>	oligonucleotide			
<400>	258 gtgt tgtgacgcaa ctgcaactcc	30		
agacocg	gege egegaegeaa eegeaaeeee	30		
<210>	259			
<211>	39			
<212>	DNA ·			
<213>	oligonucleotide			
<400> gtgatca	259 agcg gatcccttca atttagaaag caattgttc	39		
•	260			
<211>	·			
	DNA oligonucleotide			
\Z13>	oligonucleotide			
<400>	260			
	tata ttacgcacca tattc	25		
<210>	261			
<211>	22	•		
<212>	DNA			
<213>	oligonucleotide			
		22		
atacgtgacg acattggtag tc 22				
<210>	262			
<211>	15			
<212>	PRT			

PCT/GB02/00928

WO 02/070703

<213> coral

<220>

<221> misc\_feature

<222> (15)..(15)

 $\langle 223 \rangle$  x = any amino acid

<400> 262

Ser Pro Pro Asp Tyr Thr Leu Glu Phe Pro Lys Lys Xaa Val Ala 1 5 10 15

<210> 263

<211> 15

<212> PRT

<213> coral

<400> 263

Ser Pro Pro Asp Tyr Thr Leu Glu Arg Pro Lys Lys Gly Val Ala 1 5 10 15

<210> 264

<211> 24

<212> PRT

<213> coral

<400> 264

Lys Thr Asp Val Met Glu Gly Ile 20

<210> 265

<211> 22

<212> PRT

<213> coral

<400> 265

Ser Tyr Leu Pro Asn Gly Ile Ala Glu Glu Met Lys Thr Asp Leu Met 1 5 10 15

Glu Gly Ile Val Asn Gly 20

<210> 266

<211> 22

<212> PRT

<213> coral

<400> 266

Ser Leu Tyr Gln Asn Gly Ile Ala Glu Glu Met Lys Thr Asp Leu Met l $\phantom{0}$ 5  $\phantom{0}$ 10  $\phantom{0}$  15

Glu Gly Ile Val Asn Gly

<210> 267

<211> 20

<212> DNA

<213> oligonucleotide

<400> 267 atggaaggga tagtcgatgg

tggaaggga tagtcgatgg

20

20

<210> 268

<211> 20

<212> DNA

<213> oligonucleotide

<400> 268 atggaaggga ttgtcgatgg

<210> 269

234/234 <211> 20 <212> DNA <213> oligonucleotide <400> 269 20 atggaaggga tcgtcgatgg <210> 270 <211> 19 <212> DNA <213> oligonucleotide <400> 270 19 cctcgacaat cccttccat <210> 271 <211> 19 <212> DNA <213> oligonucleotide

PCT/GB02/00928

19

WO 02/070703

<400> 271

cctcgacgat cccttccat